1973

Efficient solutiosn to a multicriteria linear program, with application to an institution of higher education.

Holl, Stephen Trygve.

Johns Hopkins University

EFFICIENT SOLUTIONS TO A MULTICRITERIA LINEAR PROGRAM, WITH APPLICATION TO AN INSTITUTION OF HIGHER EDUCATION

Stephen Trygve Holl

EFFICIENT SOLUTIONS TO A MULTICRITERIA LINEAR PROGRAM, WITH APPLICATION TO AN INSTITUTION OF HIGHER EDUCATION

by

Stephen Trygve Holl

A dissertation submitted to The Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy.

Baltimore, Maryland

1973

ABSTRACT

Administrators are often confronted with problems for which there exist several distinct measures of success. Such problems can often be expressed in terms of linear programming models with several linear "criterion" functions instead of single objective functions. A variety of techniques have been applied to multicriterion problems, but there exists a need for one which does not assume technical sophistication on the part of the decision maker and which provides valid solutions with minimum effort. "Efficient Manifold Presentation," the approach used here, is based on the concept that the ideal solution to a multiobjective problem must be a Pareto optimal solution. In many cases simply narrowing the set of candidate solutions to the set of all Pareto optimal solutions may enable the decision maker to find the compromise being sought. A finite method for finding an expression for the set of all Pareto optimal solutions to a linear program with multiple linear criteria is presented. Two processes are involved; first, the discovery of all Pareto optimal vertices of the feasible region, and secondly a grouping of these into sets each of which defines a convex polyhedron of Pareto optimal possibilities. Alternate versions of the second process are suggested for use under varying circumstances. As an example of the applicability of the method, a general framework for modeling enrollment and staffing policy in an educational institution is presented, and a specific model is developed for the School of Health Services of the Johns Hopkins University.

ACKNOWLEDGEMENTS

I would especially like to thank my thesis advisor, Dr. John P. Young, without whose guidance and encouragement this work could not have been accomplished. I am deeply indebted, also, to Dr. Ramakrishna R. Vemuganti for his sharp insight and patient guidance during the development of the algorithms. Dean Peterson of the School of Health Services, and Dr. Golden and many other members of the staff of the School gave generously of their time and insight; for this I am sincerely grateful. I would also like to express my appreciation to Dr. Mandel Bellmore for sparking my interest in Pareto optimality in this context, and to all of the many members of the department of Mathematical Sciences of the Johns Hopkins University with whom I discussed this work in its formative stages. Direct or indirect credit for most of the ideas herein belong to them. I am especially grateful to Mrs. Doris Stude, who so ably typed the final copy. This work was performed while under scholarship from the Junior Line Officer Advanced Scientific Educational Program of the United States Navy; to the Navy I am eternally indebted. Finally, for tolerance and encouragement throughout, I owe an inestimable debt to my wife Kathy.

TABLE OF CONTENTS

TABLE OF ILLUSTRATIONS

[illegible]

[illegible]

CHAPTER I

INTRODUCTION

I-1 PROLOGUE: THE DEAN'S PROBLEM

A college dean is pondering the related questions of how many students his school should admit and how many faculty he should have available (or perhaps how many he should terminate) in preparation for an approaching academic year. He can predict readily enough the effects of his decisions for this upcoming year and perhaps, with somewhat more difficulty, for the year following that, but he also realizes that the potential impact of his decisions may linger long beyond. He therefore decides that he would like to address his problems in the context of several years, and after careful thought concludes that he can accurately assess the implications of five consecutive years of enrollment and staffing decisions. More specifically, he is confident that he can project numbers of students and numbers of faculty; from these he can derive a variety of revealing measurements, such as numbers of graduates, annual budget surplus, student-faculty ratios, differences between current and projected enrollments, and so on. Each such measurement is significant and none of them subsumes the others. Perhaps the Dean makes several such projections, each for a different set of enrollment and staffing assumptions. He measures each projection according

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

to those criteria he finds meaningful. It is quite likely that the criteria do not all favor the same projection. Nevertheless the Dean is able to choose a preferred projection from among the several candidates, one for which the strengths of some criteria overbalance the weaknesses of others. But is this the best possible projection? The Dean may well wonder whether he has overlooked an even more favorable possibility. He may also wonder how he would recognize the best possible projection if he were to stumble upon it.

Suppose that the Dean is able to enumerate the various criteria by which he measures alternative projections. One desirable property of this list is that it should be complete, in the sense that the information contained in the measurements according to these criteria is sufficient for him to be able to choose from among any alternatives so measured. Another desirable property is that all other things being equal, a projection with a lower "score" according to one criterion should not be preferable to one with a higher score. This should remain the case regardless of whether the scores are high or low, and regardless of how great or small the difference is between the scores. One must assume, also, that the Dean is able to specify the restrictions within which his projections must remain.

The general practice has been, and often still is, for the Dean to base his decisions upon scratch pad calculations. The general technique is frequently as follows. First he selects a set of appealing decisions, projects their impact, and scores the results according to the various criteria. Casting a critical eye over these

scores, he adjusts the decisions in a manner which he hopes will ameliorate some of the scores' deficiencies, and repeats the projection and scoring. This continues until the Dean concludes that he can do no better, or until he gets tired, or runs out of time.

If the Dean's institution has some sort of planning office, it may be possible for him to shift the burden of making the projections and computing the scores to a computer. The formulae employed to make and score the projections may constitute one of the many familiar "input-output" models. There are several ready-made computerized input-output models available, some of them extremely detailed. But regardless of the detail and precision of such a model, its focus is on projection rather than on decision. The Dean must nevertheless continue to hypothesize the decisions and judge the results. No input-output model can guarantee that the decision maker will be able to find a feasible alternative which suits him above all others, and none can guarantee that he will be able to recognize such a point should he arrive at one.

The Dean's planning office, if it is capable of more sophisticated institutional research, may suggest another approach intended to lift from the Dean the burden of decision as well as that of projection. A superficially straightforward technique consists of expressing the problem as a "mathematical program", and then solving this mathematical problem to discover the best decisions. The mathematical programming formulation might be as follows:

$$\begin{aligned} &\text{Maximize } H[f_1(x), f_2(x), \ldots, f_k(x)] \\ &\text{Subject to } g_1(x) \leqq 0 \\ &\qquad\qquad g_2(x) \leqq 0 \\ &\qquad\qquad \ldots \\ &\qquad\qquad g_m(x) \leqq 0 \end{aligned}$$

where x is a vector with components $x_1, x_2, \ldots, x_n$ representing the enrollment and staffing decisions made. The inequalities $g_i(x) \leqq 0$ express the limitations or "constraints" within which the institution operates. Each $f_i(x)$ is a function giving the score that decisions x achieve with respect to criterion i, and the function H, the problem's "objective function", numerically expresses the Dean's satisfaction with any projection measuring $f_1(x)$, $f_2(x), \ldots, f_k(x)$ with respect to the various criteria. The functions $f_i(x)$ and $g_i(x)$ are relatively easily expressed: they are, after all, the stuff of which the many input-output models are made. The satisfaction function $H[f_1(x), f_2(x), \ldots, f_k(x)]$, on the other hand, is exceedingly elusive, for it involves the comparison of incommensurables--the apples and oranges, so to speak, of the educator's trade. The Dean is able to directly compare two alternative projections and to pick the one which satisfies him most, but creating a function H which when confronted with two alternatives will unfailingly select the one which he would choose, is a very difficult task. Of course, the institutional research staff may have methods--mystical, time consuming techniques--for coaxing approximate objective functions out

of reluctant decision makers. But by now the Dean is experiencing growing discomfort about the apparently abrupt passage from grappling with abstract and perhaps esoteric satisfaction functions directly to the computer and thence to "the solution." He can hardly help but feel that he has had scant opportunity to fully employ his own expertise and intuition--that, in fact, he has now begun to lose control over the decisions for which he is responsible.

The Dean can avoid the trials of developing a comprehensive objective function and recover some of that lost initiative if he and his staff decide to attack this analysis in a different manner. Instead of formulating a mathematically complete problem and turning it over to the computer for solution, the Dean may deal with selected portions of the solution process himself, leaving the remainder to the machine. Computer solution techniques do not make use of all of the implications of an objective function anyway; loosely speaking, they merely interrogate it about the situations arising at particular alternatives. The decision maker may substitute his own judgments concerning particular alternatives for that part of the solution technique in which the computer would interrogate the objective function. The advantage of this substitution is that the objective function need never be specified precisely. Geoffrion, Dyer, and Feinberg[1] call this "interactive multicriterion optimization." A major drawback is that the type of introspection required is not as natural as the simple comparison of alternatives, for example, as in the use of an input-output model. It also involves a repetitious

cycle of introspection and computer operation, a requirement which will almost surely be inconvenient or time consuming or both. Dyer[2] has implemented interactive multicriterion optimization in a form requiring only that the decision maker be able to choose from among specific alternatives. Much indoctrination of the decision maker is avoided thereby, at the expense of increasing the computer-decision maker dialogue. Saska[3] and Benayoun and Tergny[4] provide somewhat simpler interactive optimization techniques for use when the functions $f_i(x)$ and $g_i(x)$ all happen to be linear.

Another possibility is "goal programming".[5,6,7,8] In using this approach, the decision maker sets desirable levels of accomplishment ("goals") for each criterion. Linear programming techniques may be used to minimize the total weighted underachievement of these goals, or to accomplish as many goals as possible in a specified order of importance, or some combination of these two. Setting goals and ordering priorities requires some finesse; ordinarily several successive modifications to the initial goal and priority set will need to be made.[6] The planning process using this technique is not unlike the use of an input-output model. There is, for example, no guarantee that the Dean will have the wisdom, patience, skill, or time to arrive at the most appealing possible projection.

This thesis offers another way for the Dean to approach his planning problem. The linchpin of the method is the notion of "efficiency" or "Pareto optimality," after the Italian economist who

originated the concept.[9]

Briefly, an alternative is "efficient" if it is achievable and if no other achievable alternative scores at least as well with respect to every criterion and is actually preferred with respect to at least one. In the notation presented earlier, decision set x is efficient if and only if $g_i(x) \leq 0$ for all constraints i, and if there does not exist another decision set x' such that $g_i(x') \leq 0$ for all i and such that $f_i(x') \geq f_i(x)$ for all i, and $f_i(x')$ is greater than $f_i(x)$ for at least one i.

It is obvious that if the Dean is to judge his projections solely on the basis of their scores with respect to the criteria that he has laid down, then the decision set which is most satisfying must be an efficient decision set. The method proposed is to cull the efficient decision sets from the mass of all possible decision sets and to present them, in a clear fashion, to the decision maker for his perusal. He is then free to consider these alternatives in light of their levels of attainment with respect to the various criteria laid down. The later chapters of this thesis may offer some clues to how this might be accomplished, but no firm prescriptions are possible, for balancing incommensurables is the focus of all managerial art. The signal advantage of this planning approach is that the decision maker retains control over the weighing of incommensurables in his chosen field. It is this part of the decision process for which he, above all others, is suited, and for which he alone is responsible.

The disadvantages of the method presented are that the criteria and constraints must be linear, and that sometimes the set of efficient possibilities is not very informative. The latter case can arise in a number of ways. Some problems are encountered in which every possibility is efficient; then use of the proposed method does not reduce the field of possibilities from which the decision maker must choose. It can also happen that the set of efficient possibilities, though smaller than the set of all possibilities, is still too large for the decision maker to search. Finally, even if the set of efficient possibilities is not excessively broad, so much detail may be required to express it that the decision maker may still not be able to cope with the flood of data he receives. But if one of these unfortunate eventualities should occur, the difficulty will be apparent before the Dean invests appreciable time in weighing the possibilities. Moreover, should he then turn to one of the other approaches mentioned earlier, he will probably find that he can employ the criteria and constraint expressions already developed.

I-2 MODELS FOR PLANNING IN HIGHER EDUCATION - A REVIEW OF THE LITERATURE

Most planning aids for institutions of higher education fall neatly into one of two groups. On the one hand are a host of special purpose techniques addressed to problems of daily operation, such as classroom scheduling and library operations. The range of theory employed in these tools is extensive. Schroeder[10] and McNamara[11] provide current surveys. On the other hand there are a number of broad scope institutional planning models, almost all of them of the straightforward input-output variety. Weathersby and Weinstein[12] have published a survey of such models, tabulated by their salient characteristics. These models range from simple linear forecasting of total enrollments and physical space requirements to highly involved simulations treating enrollments, faculty, financial aspects, physical facilities, and even research. The most advanced of these models are able to handle levels of aggregation, or rather disaggregation, right down to the individual lecture or laboratory session, and require data--masses of it--in corresponding detail. Apparently all models presently require their data to be specially prepared, but the creators of the most complex models--which can not be said to be finished, but are, rather, in states of continuing evolution--envision that ultimately the models will interface with pervasive institutional management information systems, thus eliminating, or perhaps camouflaging, the need for special data collections.

The two most widely known educational planning systems are CAMPUS, and RRPM, both mammoth simulations. CAMPUS was originated by Richard Judy and Jack Levine[13] at the University of Toronto. Its latest version is the most detailed educational planning model currently in operation. In technical terms, it is simply a straightforward resource-costing model, albeit a vast one. In 1970 it incorporated student flows, space and personnel needs, and support costs. Efforts have since sought to incorporate debt financing and to adapt the system to community colleges, state college and university systems, colleges with entirely individualized instruction, and primary and secondary education school systems. The model itself is resident on a large time-sharing network, and each terminal is equipped with a cathode-ray tube display in addition to normal input-output devices. The level of aggregation in each of the sectors mentioned is at the discretion of the user. In its least aggregate form, the model's data requirements are monumental; increasing the aggregation effects a corresponding reduction in the data needs. Fortunately, CAMPUS is designed to accommodate data in the forms in which it is ordinarily collected.

RRPM, the "Resource Requirements Prediction Model",[14] was originally developed by Mathematica for WICHE, the Western Interstate Commission for Higher Education. Since its inception, several versions have been developed and the RRPM models and much related work have been ceded by WICHE to NCHEMS, a new "National Center for Higher Education Management Systems at WICHE" funded by the federal

government. RRPM, too, is basically an ordinary resource-costing model. As its name implies, the principle focus of the RRPM family is upon matters of institutional finance. RRPM, however, is but one of a galaxy of projects planned or underway at NCHEMS; these include work on student and faculty dynamics, on the outputs of higher education, and on classification of higher education activities and development of precise lexicons of terms for higher education management. Each of these can be expected to have impact on future versions of RRPM. The data requirements for RRPM are compatable in format with the requirements of the Department of Health, Education, and Welfare's HEGIS ("Higher Education General Information Survey") report.

Two interesting models which are not of the ordinary input-output variety are the goal programming model of Lee and Clayton,[8] and the application of interactive nonlinear programming to the operation of an academic department, by Geoffrion, Dyer, and Feinberg.[1] Goal programming was touched upon in the last section--briefly, but in sufficient detail for present purposes. The reader will recall that goal programming is not a pure optimization technique. It is, rather, a sequence of suboptimizations dependent upon the specification of the goals. Although the procedure does proceed to optimize once goals are set and ranked, the technique does not include a formal procedure for adjusting the goals settings, weightings, and rankings. Instead, these adjustments are left to the judgment of the decision maker, just as the decisions themselves are

up to the decision maker in the case of an ordinary input-output model. In either case, progress toward an optimum is attendant upon progressive refinement of the model inputs.

On the other hand, the method of Geoffrion, et al.,[1] actually an adaptation of a well known method of solution of nonlinear programming problems, is a genuine optimization technique. It differs from the original all machine version in that the decision-maker instead of the computer is asked to interrogate the objective function. Thus, again, the decision maker is inserted into the iterative cycle. This technique will be described more fully in section I-3.

I-3 MULTICRITERION OPTIMIZATION PROBLEM ANALYSIS - A REVIEW OF THE LITERATURE

The "multicriterion optimization problem" (or "vector maximum problem") is stated[15] as follows:

$$\text{Maximize } F(x)$$
$$\text{subject to } g_i(x) \leqq 0 \quad \text{for } i = 1,2,\ldots,m$$

where x is an n-vector $(x_1,x_2,\ldots,x_n)$, and $F(x)$, a function of x, is a k-vector with scalar components $(f_1(x),f_2(x),\ldots,f_k(x))$. The solution to this problem is the set of all x which are efficient and which satisfy the constraints $g_i(x) \leqq 0$ for $i = 1,2,\ldots,m$. Solution vector x is efficient if it satisfies the constraints and if there does not exist another x' also satisfying them and having the properties that

$$f_i(x') \geqq f_i(x) \quad \text{for } i = 1,2,\ldots,k$$
$$f_i(x') > f_i(x) \quad \text{for at least one such } i.$$

Although interest in multicriterion problems is quite recent, the literature concerning these problems is already voluminous. For a bibliography and synthesis of work in this area, see Roy.[16] Roy distinguished between four approaches to the multicriterion problem, the second of which is "progressive definition of preferences together with exploration of the feasible set". In contrasting this

approach with that which begins with the construction of a unique functional, he notes substantial inefficiency in the latter. The unique functional is constructed to be able to satisfactorily rank all possibilities. This is frequently unnecessary; often some possibilities will be obviously inappropriate--it is not necessary for the functional to be able to correctly rank these. For example, it should not be necessary to deal with inefficient solutions; a proper ranking of efficient possibilities will suffice. But Roy notes that "it is often difficult to formulize the set of efficient actions and work with it". The algorithms of Chapters V and VI of this thesis offer the means to find and specify this set when the $f_l(x)$ and $g_i(x)$ are all linear expressions in the x_i.

Roy mentions goal programming as one method of "aggregation of multiple objective functions in a single function defining a complete preference order". In view of the iterative way in which the proper (or at any rate, final) goals and rankings are typically developed,[7] goal programming could just as properly be placed in Roy's second category--the "progressive definition of preferences together with the feasible set". Lee and Clayton[8] have constructed a goal programming model for a single year university planning problem, and Lee and Sevebick[7] have applied goal programming to a dynamic, multiple year municiple planning problem. Both of these models are entirely linear. Dyer[17] has linked goal programming with the more formal "interactive" exploration of the feasible region and objective function approach of Geoffrion, Dyer, and Feinberg.

In the latter, a proven nonlinear programming technique is employed in a manner requiring only "local" information about the decision maker's preferences. Essentially, Geoffrion et al.[1] suggest the following iterative cycle. Beginning at some particular solution, the decision maker is asked to state tradeoffs ratios among the criteria evaluated at that point. This information is translated into a direction of improvement; the computer then determines the criteria values achieved as the solution is varied along the direction chosen. These results are displayed for the decision maker, who selects the distance along that direction which seems to achieve the most desirable results. At the solution so selected, the decision maker states new criteria tradeoffs, and the process continues as before. An optimum is found when there is no direction of improvement.

Roy cites two other interactive approaches,[3,4] both involving an iterative interrogation/computation cycle, and both of them functioning in a linear programming framework. The technique of Geoffrion et al.[1] is, of course, not restricted to the linear case.

The characterization of efficiency in the mathematical programming context is admirably summarized and extended in Philip;[18] it is on this paper that much of the following analytic work is based. The most useful result in the paper is an extension of a theorem of Kuhn and Tucker providing useful necessary and sufficient conditions for efficiency. The theorem is that if x is on the boundary of a polyhedral feasible region then x is efficient if and only if some

nonnegative weighting of the gradients of the constraints binding at $\ddot{x}$ equals some weighting of the gradients of all of the objective criteria, where these (latter) weights are all greater than or equal to 1. This condition can be conveniently checked by attempting a linear program. Much use of this will be made in Chapters V and VI. Philip also presents a collection of algorithms, some due to others and some his own, which are addressed to the problems of deciding if a solution is efficient, finding an efficient solution, finding if a given efficient solution is unique and if not finding another, and finding an efficient solution which maximizes some linear objective function. He does not address the problem of finding and expressing the set of all efficient solutions.

Finally, some interesting work is being done in the area of dominance structures.[19] Efficiency is the weakest possible form of dominance; the strongest is supplied by the usual single objective function. Much remains to be done to enable decision maker and analyst to deal effectively with problems for which there is readily available a dominance structure somewhere between these extremes.

I-4 THE CASE FOR EFFICIENT MANIFOLD PRESENTATION

The approach to multicriterion optimization problems which is suggested in this thesis consists of presenting to the decision maker the "efficient manifold"--that is, the set of all efficient, or "Pareto optimal", possibilities. For convenience, this method will be styled "efficient manifold presentation", or simply "EMP". The case for using efficient manifold presentation rather than an input-output model, goal programming, or straightforward or interactive nonlinear programming revolves primarily around its improved accommodation to the difficult circumstances of the average decision maker, to wit:

a) Instead of offering "the answer", it provides a range of possibilities over which the decision maker can exercise his insight, taking into account unquantifiable and intangible or simply unrecognized considerations. The absence of choice in "the answer" is coming to be recognized as a major failing in much of management science's mathematical modeling.[20]

b) It allows the decision maker to apply his managerial arts and his intuition in a familiar context--that is, directly to the problem of choosing the best alternative from among a set of real alternatives. Most other approaches permit only indirect influence over the course of the method, for example through parameter adjustments, goals setting and ranking, or formulation of objective

functions. Any of the latter is apt to seem unnatural to the decision maker, and each is required before or during use of the computer, if one is employed, rather than after.

c) EMP does not require the decision maker to interact with the computer. Besides being inherently distasteful to the uninitiate, interaction either requires the decision maker to be at a computer terminal whenever he works on his problem (most executives have neither the time nor the inclination to use computer terminals), or else it imposes a considerable time lag while the computations for each new iteration are batch processed.

d) Unlike the input-output model approach, but in common with the other approaches, EMP fosters thought about criteria as well as constraints. This enforced discipline can by itself be expected to be very enlightening, even if EMP is not ultimately pursued.

e) Most importantly, EMP, if successfully employed, can economize on the decision maker's most precious commodity --his own time.[21] Because it requires neither iterative computer interaction nor difficult and time consuming definition of a monolithic preference function, it is anticipated that EMP will prove to be a most streamlined aid for multicriterion decisionmaking.

Of course, EMP is no panacea. The theory in this thesis only enables one to cope with problems whose variables are continuous, whose constraints are linear, and whose criteria are linear and monotonically desirable. In addition, if the number of criteria is excessively large, or if the set of efficient possibilities is very complex, then the EMP method may generate more information than the decision maker can digest. Finally, in some problems every possibility is an efficient possibility. In such a case, the EMP method will result in explicit expression of the set of all possibilities and if there are any constraints which are nowhere binding, this fact will be exposed. (Such a case arises in the example problem of Chapter IV and VII.) So some enlightenment is possible even here.

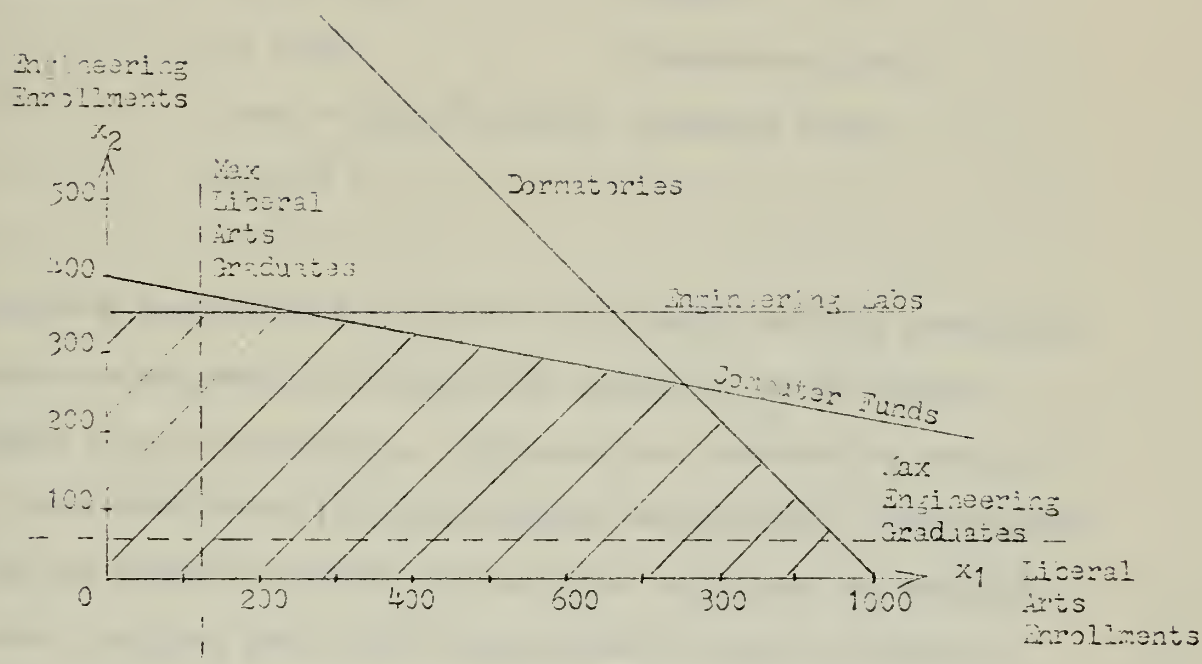
[illegible]
[illegible]
[illegible]
[illegible]
[illegible]
[illegible]
[illegible]
[illegible]
[illegible]

[illegible]
[illegible]
[illegible]

I-5 A SIMPLE EXAMPLE OF EFFICIENT MANIFOLD PRESENTATION

As a rather unrealistically simple example of the use of efficient manifold presentation, consider again the case of the Dean of a small college who is deciding upon a target steady state enrollment for his school. His institution produces two types of graduates; liberal arts graduates (enrollment in the liberal arts curriculum will be represented by the variable x_1), and engineers (enrollments in the engineering program will be x_2). The annual number of students who graduate from each program is estimated to be 20% of the total steady state enrollment in that curriculum. The Dean is not indifferent to the mix of graduates; in particular, he feels that a heavy preponderance of one type of graduate would be less desirable than an equal number of students more evenly distributed. Consequently he has two criteria; the yearly number of liberal arts students graduated, and the yearly number of engineers graduated.

Three constraints limit enrollments at the Dean's college. First, all students must live in the dormatories, and dormatory accommodations--and therefore total enrollments--are limited to 1000. Engineering laboratory space is also limited; no more than 350 engineers can be accommodated. Finally, the institution's annual computer budget is a firm $200,000. Usage rates at the institution seem to average $100 per liberal arts student per year, and $500 per engineer per year.



Engineering
Enrollments
x_2
500
400
300
200
100
0
Max
Liberal
Arts
Graduates
Dormatories
Engineering Labs
Computer Funds
Max
Engineering
Graduates
200
400
600
800
1000
x_1
Liberal
Arts
Enrollments

Figure I

The Dean's problem is depicted graphically in Figure I and algebraically below. The total enrollments corresponding to each point in the graph may be read off of the axes; the graduation rates are 1/5 of these.

Criteria:	Max $1/5\ x_1$	(max liberal arts graduates)
	Max $1/5\ x_2$	(max engineering graduates)
Constraints:	$x_1+x_2 \leqq 1000$	(dorms)
	$x_2 \leqq 350$	(engineering labs)
	$100x_1 + 500x_2 \leqq 200{,}000$	(computer funds)
	$x_1, x_2 \geqq 0$	

The three constraints are plotted in the graph; the only enrollments which are achievable are those whose points lie in the hatched region or on its boundaries. The graph also contains two arbitrarily positioned dotted lines representing the criteria. Loosely speaking, the object is to push the horizontal one up and the vertical one to the right until their intersection is about to leave the hatched region. <u>The point so found will be only one of a number of possible points with the property that further improvement of either criterion is impossible except at the expense of the other.</u>

The Dean's institutional researcher wishes to present to the Dean the set of all efficient enrollments. The observant reader will recognize that an efficient point must lie in or on the boundary of the hatched region and must have the property that there is no

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

other point in or on the boundary of the hatched region which scores at least as well as this one with respect to both criteria and scores better with respect to at least one of the criteria.

These efficient points are the ones lying on the heavy lines in Figure II. The reader should verify that for any other point in the hatched region, there exists another which is at least as far upward and to the right, and further in one or both directions. For example, of the five lettered points in this graph, only two--D and E--are efficient. Point A is not efficient because other achievable points, including E, have greater numbers of both types of graduates; similarly, B is "dominated" in like manner by point D. There are no achievable points with more of both types of graduates than are achieved at C, but there are some, including D, with at least as many engineers and more liberal arts graduates; then C is not efficient either.

The Dean is now able to restrict his attention to the possibilities represented by the heavy lines, rather than having to consider all of the possibilities in the hatched region. After some deliberation, he chooses the enrollments set represented by the point E, because he feels most comfortable with the numbers of graduates of the two types which it offers. He does this without considering any alternatives other than those on the heavy lines, but this is sufficient because to every alternative in the hatched region but not on the heavy lines there corresponds another on the heavy lines which is clearly preferable to the first. Since point E

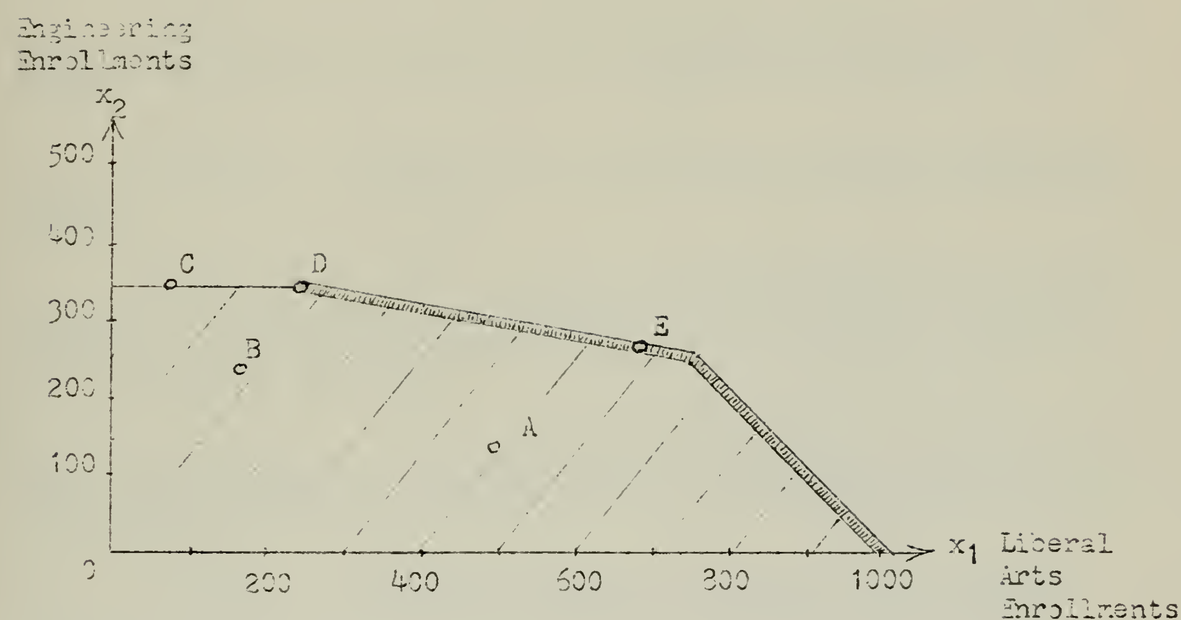

Engineering
Enrollments
x_2
500
400
300
200
100
0
C
D
B
E
A
200
400
600
800
1000
x_1
Liberal
Arts
Enrollments

Figure II

was preferred to this second point, it must also be preferred to the first.

FOOTNOTES

[1]Geoffrion, A. M., J. S. Dyer, and A. Feinberg, "An Iterative Approach for Multi-Criterion Optimization with an Application to the Operation of an Academic Department," Management Science, XIX (1972), Applications series, pp. 357-368.

[2]Dyer, J. S. "A Time-Sharing Computer Program for the Solution of the Multiple Criteria Problem," Management Science, XIX (1973), Applications series, pp. 1379-1383.

[3]Saska, J. "Linear Multiprogramming," Economiko Matematiky Obzor, IV (1968), pp. 359-373.

[4]Benayoun, R. and J. Tergny, "Critères multiples en programmation mathematique: une solution dans le cas lineaire.", Revue Française d'Informatique et de Recherche Operationnelle, 3ème année, no. V-2 (1969), pp. 31-56.

[5]Charnes, A. and W. Cooper, Management Models and Industrial Applications of Linear Programming, Vol. I, John Wiley and Sons, New York, 1961.

[6]Ijiri, Y. Management Goals and Accounting for Control, North-Holland Publishing Company, Amsterdam, 1965.

[7]Lee, S. M. and W. Sevebick, "An Aggregation Model for Municipal Economic Planning," Policy Sciences, II (1971), pp. 99-115.

[8]Lee, S. M. and E. R. Clayton, "A Goal Programming Model for Academic Resource Allocation," Management Science, XVIII (1972), Applications series, pp. B395-B408.

[9]Miller, D. W. and M. K. Starr, The Structure of Human Decisions, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1967, p. 56.

[10]Schroeder, R. G. "A Survey of Management Science in University Operations," Management Science, XIX (1973), Applications series, pp. 895-906.

[11]McNamara, J. F. "Mathematical Programming Applications in Educational Planning," Socio-Economic Planning Sciences, VII (1973), pp. 19-35.

[12]Weathersby, G. B. and M. C. Weinstein, A Structural Comparison of Analytic Models for University Planning, University of California, Research Program in University Administration, paper P-12, August 1970.

[13]Judy, R. W., J. B. Levine, and S. I. Center, Campus V Documentation, Vols. 1-6, Systems Research Group, Toronto, Canada, 1970.

[14]Resource Requirements Prediction Model (RRPM) 1.6; Users Manual, National Center for Higher Education Management Systems at WICHE, Boulder, Colorado, 1972.

[15]Geoffrion, A. M. "Proper Efficiency and the Theory of Vector Maximization," Journal of Mathematical Analysis and Applications, 22 (1968), pp. 618-630.

[16]Roy, B. "Problems and Methods with Multiple Objective Functions," Mathematical Programming, VI (1971), pp. 239-266.

[17]Dyer, J. S. "Interactive Goal Programming Management Science, XIX (1972), Applications series, pp. 62-70.

[18]Philip, J. "Algorithms for the Vector Maximization Problem," Mathematical Programming, II (1972), pp. 207-229.

[19]Yu, P. L. Cone Convexity, Cone Extreme Points and Nondominated Solutions in Decision Problems with Multiobjectives, University of Rochester, Center for System Science paper CSS 72-02, April 1972.

[20]Vazsonyi, A. "Why Should the Management Scientist Grapple with Information Systems," Interfaces, III, Number 2, pp. 1-18.

[21]Mintyberg, H. "Managerial Work: Analysis from Observation Management Sciences, XVIII (1971), Applications series, pp. B97-B110.

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

CHAPTER II

THE SCHOOL OF HEALTH SERVICES

II-1 RATIONALIZATION AND REALLIGNMENT OF THE NON-PHYSICIAN HEALTH CARE PROFESSIONS: A NEED AND AN OPPORTUNITY

It is common knowledge that health care delivery is a troubled industry today. Though it is not the purpose of this thesis to air the maladies of the medical world, certain by no means original observations are in order. Recent decades have seen the specialization of almost all physicians; this trend has left a huge and growing gap in primary health care. The breach could be filled by non-physician intermediate-level health care professionals, were they numerous enough, but unfortunately they are not. In fact, there is a serious shortage of persons with baccalaureate level training in health care delivery. Nurses currently comprise the largest proportion of such professionals and nurses are indeed in short supply, but the most critical shortages are in the less well known, more recently established and more quickly growing professions.* It is expected that much of the future unfilled demand

---

*The breadth of this family of health care professionals is illustrated by the list of curricula for the proposed School of Health Services of the University of Nevada.[1]

Anatomy	Medicine
Biochemistry	Nursing

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

for baccalaureate level health care professionals will be for members of these new fields, and for professionals whose professions have not yet been defined. Unfortunately, the newest professions are unsupported by educational establishments equal to the task of training sufficient numbers of graduates in the new fields, and even the long established schools of nursing are beset both by financial problems and by a crumbling of the general consensus on what nursing education should encompass. Educational institutions which might be interested in contributing to the solution are handicapped by a lack of agreement on the character of the new professions and on the forms of the curricula that students of these professions should follow. Compounding this, the new depression in higher education, with its general decreases in state and federal funding support, discourages experimental ventures.

A complicating factor in the shortage of intermediate level health care professionals is the absence of well defined paths for advancement. Provision for career ladders for these professionals is a very difficult problem, one which will require considerable study before appropriate solutions are found. Unfortunately,

---

Biomedical Engineering	Pathology
Biomedical Instrumentation	Physical Therapy
Dietetics	Physiology
Health Education	Science Writing
Medical Technology	Speech Pathology

Even this extensive list is not exhaustive.

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

solutions to this problem and to the problems of curricula organization are intimately related.

New curricula, facilities, and agreement on professional boundaries are needed to alleviate this manpower shortage. Because of the decentralized nature of higher education, these are apt to develop slowly. Progress will be most rapid if the more prestigeous institutions of medical education will take the lead. As a recognized pace setter in medical education, it is fitting that the Johns Hopkins University should pioneer in these advances. It is the recognition of these needs and this mandate which led this University to create the School of Health Services.

II-2 GENESIS OF THE SCHOOL OF HEALTH SERVICES

The Johns Hopkins University and Hospital have long participated in education for the allied health professions, both through a distinguished diploma program in nursing and through numerous _ad hoc_ programs for training intermediate and lower level health care personnel. Naturally, as in all institutions of higher education, the organization and content of curricula have been under constant study and discussion by a host of committees and individuals. The chain of events leading to the establishment of the School of Health Services, heir to this long succession of interest and participation in non-physician medical education, may be said to have formally begun in 1967 when an _ad hoc_ committee of the Faculty of Medicine under the chairmanship of Dr. Russell Morgan recommended the formation of a College of Allied Medical Sciences. This recommendation was approved by the Advisory Board of the School of Medicine and in 1968 the means of implementation became the subject of study of a University-wide committee chaired by Dr. Morgan. As a member of this committee, Dr. John P. Young, Associate Provost for Planning and Advanced Policy Studies, prepared an extensive and detailed feasibility study. The study, which was completed in December of 1968, proposed the initial establishment of two baccalaureate programs, one in nursing sciences and another in clinical laboratory sciences, and included estimated costs and curricula outlines. This report was endorsed by the Morgan committee and provided the focus for further discussion and deliberation, in 1969, in two informal joint meetings of the Advisory Boards of Medicine and Hygiene, and

[illegible]

the Academic Council of the Faculty of Arts and Sciences, during which the nascent division was given approval in principle. In December of 1970, the Center for Allied Health Careers was established under the direction of Dr. Dennis G. Carlson and charged with continuing to plan for the education of lower and middle level health care professionals. The Center produced A Planning Report for Education and Training in Health Services in May of 1971 (revised in July of the same year) which again called for the establishment of a distinct school within the University, under the title "School of Health Services". In another significant development that summer the Johns Hopkins Hospital, which is organizationally distinct from the University, commenced the closure of its three year diploma program in nursing by not admitting students for the class of 1974. In October of 1971 the Board of Trustees of the University approved the formation of the School of Health Services and on 1 March 1972, Dr. Malcolm L. Peterson was appointed Dean of the new school. In February of 1973 the Board of Trustees authorized the appointment of a faculty and admission of students for the following academic year.

II-3 CHARACTERISTICS AND CURRICULA OF THE SCHOOL OF HEALTH SERVICES

The new School of Health Services fell heir to several health care personnel education programs which existed prior to the School's formation. These included an associate degree program for Health Assistants, operated in collaboration with Essex Community College, and a course of instruction for Pediatric Nurse Practitioners, operated in collaboration with the Department of Pediatrics of the School of Medicine.

The first new program to be initiated by the new School will be the Health Associates program, which will prepare individuals to deliver health care in ambulatory settings. The curriculum is to be two years in length; incoming students will be required to have completed at least two years of college before matriculating. At the end of the course, a baccalaureate degree will be awarded. The curriculum for this program will be presented in a unique format. Instead of courses organized by traditional disciplines (biology, chemistry, etc.) material will be presented in "modules" directed toward commonly encountered health care problems. The typical module will contain elements of instruction in several branches of theory and in pertinent tasks and skills. Faculty who will be involved in the Health Associates program can be divided into three broad categories. A "primary teaching team" will be responsible for most of the instruction of each group or "section" of 25 students. The members of the primary teaching team will be one physician, two "social generalists" (anthropologists, mental health counselors,

social workers, health educator/community organizers, etc.), and three "health practitioners" (persons with approximately the qualifications of the graduates of this or the nurse practitioner program to be described later). Supplementing the primary teaching team will be faculty specialists (social scientists, surgeons, pathologists, etc.) and clinical supervisors (who will oversee practical training on location in various health care delivery settings). Finally, the faculty will be supported by an extensive media resources program whose function is to amplify their effectiveness through the employment of modern technical educational aids.

The second baccalaureate program planned is the Nurse Practitioner program. It is expected that this program will be very similar to the Health Associates program except that its focus will be on health care delivery in institutional care settings rather than in ambulatory settings. Other baccalaureate programs under consideration for the immediate future are programs in Health Services Management, Environmental Hygiene, and Clinical Laboratory Science.

Additional new programs are planned for the more distant future. It is anticipated that an undergraduate program in midwifery and an Adult Nurse Practitioner program similar to the ongoing Pediatric Nurse Practitioner program may be initiated. Also planned are a program in Nuclear Medicine Technology, and one in Clinical Laboratory Technology to be conducted jointly with Essex Community College in the same manner that the Health Assistant

[illegible]

program is now operating.

While much study of the role, mission, and feasibility of the new School was done prior to its establishment, virtually all of the decisions relating to program and curricula, staffing and admissions, funding, and accession and allocation of physical facilities were left to the staff of the School. The urgent need for the School's graduates and the undesirability of a long preparatory period before tuition and operating subsidy grants could be realized dictated that the School commence operations as soon as practicable; accordingly, it was decided at the onset to focus on development of the curriculum for the Health Associates program, the first new program to be instituted. Because both the Health Associate concept and the modular teaching format are innovations, development of curriculum is necessarily slow and laborious. In addition to this curriculum development, the administrators of the School must locate faculty and funding, decide upon and implement the mechanics of admissions, and acquire facilities and equipment. All of these tasks will continue as long as new programs are being introduced (at the rate of one baccalaureate program annually, at least through 1977). Concurrent with all of these immediate responsibilities is the less urgent but nevertheless important problem of longer range planning--particularly the planning of relative emphasis on curricula and of projected requirements for personnel, facilities, and funding. Under the circumstances, this planning is exceptionally difficult to carry out. Unlike an ongoing institution, the School of Health

[illegible]

Services has no historical basis from which to extrapolate into the future, no current faculty or students to influence ranges of choice, no traditional commitments for facility or financial support. Moreover, because the very curricula are new and the mode of instruction is unprecedented, only the most general sorts of inferences can be drawn from the experiences of other institutions. Also, curricula formalize and specify the relationships between students and staff, so that at least this important information will not be known with certainty until all curricula are developed--an eventuality which is years in the future. Nevertheless, decisions will be made in the interim, if only by default, and they can only be made with the imperfect information available.

One set of longer range decisions to be made involves enrollments and staffing levels. In particular, the distribution of the School's limited resources over the various baccalaureate programs must be made. A number of planning approaches potentially useful in higher education were listed in Chapter I; of these, efficient manifold presentation is perhaps the one best suited to these particular circumstances, for the following reasons. To begin with, the relatively small size of the School makes complex computerized input-output models a poor choice. Since the staff of the School is working under the press of many immediate problems, the computer-interactive approaches may be ruled out as too time consuming. Similarly, there is scant opportunity for the difficult development of a composite objective function for use in ordinary

nonlinear programming optimization. Efficient manifold presentation, on the other hand, suffers none of these drawbacks. It has the potential to provide significantly better insight than scratch pad input-output projections for very little additional decision maker effort.

The next chapter contains an assortment of components suitable for constructing linear multicriterion educational planning models--that is, planning models for higher education which are amenable to efficient manifold presentation. The following chapter contains a model tailored to the circumstances of the School of Health Services. The solution technique is developed in Chapters V and VI, and the results of its application to the School of Health Services model are given in Chapter VII.

FOOTNOTES

[1] Smith, G. T. and N. A. Baxter, University of Nevada Medical Education Feasibility Study, Nevada State Bureau of Business and Economic Research, November 1968.

CHAPTER III

A FRAMEWORK FOR ENROLLMENT AND FACULTY PLANNING

III-1 GENERAL

This chapter contains a general model of certain aspects of the operation of an educational institution and a broad brush characterization of the use of the "efficient manifold presentation method" to augment the intuition when considering admissions and staffing decisions. In Chapter IV, this model will be tailored to fit the circumstances of the School of Health Services. As an illustration of the use of the efficient manifold presentation method, the results of its application to the model of Chapter IV are presented in Chapter VII.

It is natural to think of planning problems in terms of (1) ends or goals, and (2) circumstances prescribing the allowable actions which could be employed in the pursuit of these goals. The reader will recall from Chapter I that the actions of interest here are student admissions and faculty staffing decisions. These, the "variables" for the following model, will be represented by the symbols c and f, appropriately subscripted. The goals and circumstances affecting or affected by these variables are the criteria and constraints. In the following two sections, some criteria and constraints which are likely to arise in admissions and staffing planning for institutions of higher education are discussed.

III-2 SAMPLE CRITERIA

From the modeler's point of view it is quite unfortunate that of the many advertised goals of the world of higher education, few seem to be susceptible to precise definition or accurate measurement. Nevertheless, the following list of candidate criteria is offered. This does not pretend to be a comprehensive list of all the motives that propel educators; it does purport to be a reasonably complete list of the measurable goals whose achievements are noticeably affected by decisions on numbers of admissions and faculty. The criteria are as follows:

Graduate as many high quality students from a given program as possible (applicable to each degree program)

Maximize the excess of income over expenditures--or minimize the shortfall

Minimize the size of the average class in a given category of courses (applicable to each course category)

Minimize the deviation from a specified ideal discipline distribution within the faculty

Minimize the deviation from a specified ideal curricular distribution within the student body

Minimize faculty terminations

Minimize the additional capital investment required to support the chosen enrollment-staffing plan

Before discussing each of these objectives individually, some general remarks are in order. First, none of these criteria is

intended to represent, by itself, a capsule philosophy of management for any educator. It must be borne in mind that these criteria are members of a family of yardsticks with which the thoughtful administrator measures the performance or anticipated performance of his, and other, institutions. None of these criteria claims to provide any ordering relationship by itself except when all other criteria are equal. As the explanations of these criteria and of the constraints in the next section unfold, the opportunity will be taken to introduce equations and some occasional notation, in anticipation of the formal mathematical model of Chapter IV. Some of the criteria above are stated as quantities to be maximized; others as quantities to be minimized. For uniformity, and to be consistent with the theory developed later, all of the mathematical expressions for these criteria will be expressed as quantities to be maximized.

Graduation rates:

Heading the list of criteria is that of graduating as many high quality students as possible from a given program. The virtue of this--which used to be self-evident--is somewhat tarnished by the present surfeit of graduates in a number of fields and by the perennial glut in a few; nevertheless this is a plausible criterion if only because the just, equitable, and economical way to limit output is to put additional emphasis on other worthy and partially conflicting criteria, such as annual surplus. Additionally, when all other things are equal, it is almost certainly the case that in the eyes

[illegible]

[illegible]

[illegible]

of any particular institution, more of its own graduates is preferred to fewer, even when the market is flooded.

The mathematical quantity to maximize is the product of enrollment in the program, or if enrollments are categorized by program year then the enrollment in the last year of the program, times a constant representing the proportion of the enrollment who become graduates each year.

Annual financial surplus:

In the literature of institutional research, one occasionally finds the university described not as a "non-profit" institution but as a "not-for-profit" institution. The distinction is more important and less subtle than it appears. The days of profit oriented proprietary schools are indeed gone, but the need for fiscal responsibility is not. Financial surplusses are still desirable, for they can be devoted to any number of desirable uses, from improving quality to increasing endowment. The second criterion is to maximize profits; this enables one to distinguish between enrollment-staffing plans which are indistinguishable except with regard to expense. The mathematical quantity to be maximized is the sum over all students of the net income per student less the sum over all faculty of the net expense per faculty member less a constant subsuming all expenses and incomes not proportional to students or faculty.

Maximizing the sum excluding this last constant amounts to

the same thing. Significance is to be found in the relative rather than absolute values of the scores according to this criterion.

Student/faculty ratios:

The third criterion is to minimize average class sizes. The belief that knowledge is better transmitted to small groups than to large groups is a prominent feature of educational lore; it is expected that when all other things are equal, enrollment-staffing plans with small classes are preferred. The natural way to express the relationship between students, faculty, and class sizes would be the following:

$$\text{average class size} = \frac{(\text{constant})\text{x}(\text{numbers of students})}{(\text{numbers of faculty})}$$

where the constant subsumes average course loads and faculty teaching loads and the like. Unfortunately this equation is not linear in the two variables, and linearity is one property which is essential for the analysis of Chapter V. One way to deal with this problem is to use two separate criteria: numbers of faculty (to be maximized), and the negative of the numbers of students (to be maximized).

Faculty composition:

The composition of the faculty by disciplines is important for a number of reasons. The most obvious is that the faculty in

each discipline should be numerous enough to teach the courses in that discipline which are needed to meet the requirements of the students enrolled. This requirement will be expressed as a constraint in the next section. Another family of constraints could arise from considerations, such as accreditation standards, requiring at least a certain proportion of the faculty to hold particular types of degrees. In addition to these, the disciplinary composition of the faculty is important in its own right. For example, virtually every institution will insist on maintaining a humanities program even if it is not required by one of the constraints just mentioned. The reason is the humanizing effect these faculty are expected to have on students and on their own more technical colleagues. If the Dean is able to express his notion of the ideal community of scholars in terms of percentage composition, then the deviation from that ideal can be expressed and minimized. Suppose that it is desired that each faculty type should comprise the proportion b_i of the entire faculty. Introduce the following constraints, one for each category i:

$$b_i \sum_i f_i - f_i + x_i^- - x_i^+ = 0$$

where f_i is the number of faculty of category i, and x_i^+ and x_i^- are the positive and negative differences between the actual and desired proportions, and where all of these variables are constrained to be nonnegative. Then the quantity to be maximized is:

[illegible]

[illegible]

[illegible]

$$\text{Maximize} \sum_i -x_i^+ - x_i^-$$

Involuntary faculty terminations:

One of the most difficult, painful, and controversial activities which circumstances sometimes force a Dean to carry out is the task of reducing the size of the faculty. A most reasonable objective is to reduce as much as possible the expected number of involuntary (nonretirement) terminations which are built into any given enrollment and staffing plan. The number of faculty of category i in a given year k can be approximated by the following equation:

$$f_{i,k} = \mu \sum_i f_{i\ k-1} + h_{ik} - r_{ik}$$

where f_{ik} is the number of faculty of type i in year k, μ is the proportion of a previous year's faculty who remain voluntarily or wish to do so for the year in question, h_{ik} is the number hired, and r_{ik} is the number whose terminations begin with the year k. If this constraint is added for each faculty category, then the quantity to be maximized is:

$$\text{Maximize} \sum_i -r_{ik}$$

Capital investment:

The last criterion on the list is to minimize the amount of

$$[illegible]$$

[illegible]

[illegible]

$$[illegible]$$

[illegible]

$$[illegible]$$

[illegible]

[illegible]

additional capital investment (e.g. new construction) required. If α_{ik} and β_{jk} are the amounts of facility type k required per faculty member of category i and per student of variety j respectively; if E_k is the amount of this facility already available; and if d_k is the capital investment per additional unit of this facility, then adding the following constraint for each k

$$\sum_i \alpha_{ik} f_i + \sum_j \beta_{jk} c_j - x_k \leq E_k$$

$$(x_k = 0)$$

allows use of the following criterion:

$$\text{Maximize} \sum_k -d_k x_k$$

[illegible]

[illegible]

$$[illegible]$$

[illegible]

[illegible]

[illegible]

[illegible]

III-3 SAMPLE CONSTRAINTS

A few constraints have already been mentioned in conjunction with the criteria of the last section. Some additional constraints are:

Enrollments, numbers of faculty, and student-faculty ratios are related through a mathematical identity.

Admissions and enrollments are mathematically related.

The number of applications sets a limit to the number who ultimately matriculate.

The number of prospective faculty available limits the amount of hiring which may be done.

The facilities available limit enrollments and staff.

Upper (lower) bounds on the proportion of the faculty satisfying certain characteristics may be imposed.

Upper (lower) limits on the proportion of enrollments composed of certain varieties of student may be imposed.

Student/faculty proportionality:

The first two constraint types are actually mathematical identities expressing obvious but occasionally necessary relationships. In the last section it was noted that student-faculty ratios are of great importance, and a pair of criteria based on minimizing the class size was proposed. In the event that it is desired to fix the class size, then the following constraint is in order:

$$\#\text{faculty} = \frac{(\#\text{students})\text{x}(\text{rate of course registrations per student})}{(\text{class stu-fac ratio})\text{x}(\#\text{courses per professor})}$$

The only variables here are numbers of students and faculty, so this is linear. When faculty is classified according to discipline and students according to curriculum or according to curriculum and year of curriculum then the last equation becomes a family of equations, one for each faculty category. The matrix of course registrations by discipline and curriculum, or by curriculum and year of curriculum, is called the "induced course load matrix". This "ICLM" is an almost ubiquitous feature of higher education planning models; it is frequently the Achilles heel of the model, too, because when the curricula have room for electives, the entries in the matrix depend on the mercurial interests of the student body. That is, when the curricula are flexible, the ICLM is difficult to establish and prone to surreptitious change.

Relationships among matriculations and enrollments:

The second constraint may be expressed:

$$\begin{array}{l} \text{number of} \\ \text{students} \\ \text{next year} \end{array} = (\text{number this year}) \text{ x } \left(\begin{array}{l}\text{proportion} \\ \text{who stay}\end{array}\right) + (\text{new students})$$

This constraint becomes a family of constraints when the student body is categorized. The proportionality figures here will be knowable to a greater or lesser extent depending on whether curricula

guidelines are firm or flexible and whether it is difficult or easy for the student to transfer from one curriculum to another.

Staff and student availability:

The next two constraints say simply that no more students (faculty) may be matriculated (hired) than are available. Some authorities forecast a decline in college attendance in the coming decade, so matriculations constraints are apt to be realities for many institutions. Despite the very visible abundance of advanced degrees in some fields, genuine faculty availability limits will not be unknown either. Such a case arises in the example of the next chapter. If p_{ik} is the number of type i students matriculated in year k and P_{ik} is the number of bona fide applicants who are acceptable that year, then the expression is

$$p_{ik} \leq P_{ik}$$

The analogous faculty inequality is

$$h_{ik} \leq H_{ik}$$

where h_{ik} is the number of category i faculty hired in year k and H_{ik} is the number of such faculty available then.

[illegible]

[illegible]

[illegible]

$$[illegible]$$

[illegible]

$$[illegible]$$

[illegible]

Facilities:

Associated with any operation of the size and complexity of an institution of higher education will be myriads of physical restrictions that the operation must stay within. Most of these never prove a hindrance because other restrictions take effect first or because the desired course doesn't happen to pass anywhere near this particular shoal--but it is not always possible to predict which restrictions are going to prove to be real, and which are not, so it is advisable to add the not unlikely ones to the model just in case. Some of the more likely to be encountered are numbers of classrooms, of laboratories, and of offices. Medical curricula have special requirements and, correspondingly, special constraints such as limited numbers of patients, wards, and examining rooms. The standard expression for such a constraint is:

availability $\geq$ sum over all users of each user's demand for the facility

Upper and lower bounds on types of faculty:

Bounds on certain components of the faculty might arise, for example, from accreditation requirements. Whatever the motivation, the appropriate expression for an upper bound is:

$$\sum_i (g_i-G)f_i \leq 0$$

where g_i is the proportion of faculty of category i who have the characteristic of interest, and G is the specified upper bound on the proportion of the entire faculty who share this characteristic. In the case of a lower bound, the sense of the inequality is reversed.

Upper and Lower bounds on the curricular composition of the student body:

This last type of constraint is of exactly the same form as that just discussed. An upper bound constraint is:

$$\sum_i (q_i - Q)C_i \leq 0$$

where Q_i is the enrollment in curriculum i, and q and Q refer to students in a manner analagous to g and G above. For a lower bound, the following constraint is in order:

$$\sum_i (q_i - Q)C_i \geq 0$$

III-4 A CHARACTERIZATION OF THE EFFICIENT MANIFOLD

The uses and limitations of the "efficient manifold presentation" method were summarized in Chapter I, section 4. In this section, the form of the efficient manifold will be delineated, a method for finding and characterizing it described, and the relation between the form of the efficient manifold and the character of the model from which it derives will be discussed. Some of the language of linear programming will be employed, but rigor is postponed until Chapter V.

Imagine that the Dean and his staff have formed a model of their institution, cobbled out of the raw materials presented in sections 2 and 3 of this chapter, with perhaps some additional criteria or constraints to fit the particular circumstances of their institution. This model is entirely linear. The question they must now address is how to find the set of all efficient solutions, the "efficient manifold."

Recall the definition of an efficient solution; it is a solution to the constraint equations which has the property that there exists no other solution to those equations (in the linear programming vernacular, no other "feasible" solution) which is at least as desirable as this one with respect to each separate criterion and more desirable according to at least one. Now suppose that all of the criteria in the model are added together to form one objective, and let us suppose that the linear program with this objective and with the constraints of the Dean's model is solved. If there are

solutions but no finite optima, then the model is badly formulated, for no real world problem with physical interpretations can have an unlimited optima. If there are no solutions to the constraint set, then either the model is poorly formulated or the situation of the institution is hopeless. If there is a finite optimum, then this optimal solution is efficient. For if it were not, there would exist another feasible solution with the property that it scores at least as well with respect to each criterion and better with respect to at least one. Adding these scores together gives a score for the linear program objective which is greater than that of the supposed optima, which is of course impossible.

This first efficient solution happens to be a vertex of the feasible region--a "corner" on the polyhedron of possible alternatives, because it is a basic solution to the set of constraint equations.

A consequence of a theorem in Chapter V is that either the efficient manifold comprises the entire feasible region or else the efficient manifold is contained entirely in the bounding hyperplanes (surfaces of any dimension less than the dimension of the problem space) of the feasible region. More than that, the efficient manifold can be expressed as a collection of polyhedra formed by the intersection of the feasible region with supporting hyperplanes (thus if the feasible region is three dimensional, its efficient manifold--presuming that it is not the entire feasible region--is some collection of the faces, edges, and vertices of the feasible

region). Incidently, every corner of such a polyhedron formed by the intersection of the feasible region and a supporting hyperplane is also a corner of the feasible region; thus the intersection can be represented as the set of all convex combinations of some set of efficient vertices of the feasible region.

Another consequence of the theory of Chapter V is that the efficient manifold for such a problem must be connected. From this it is clear that any efficient vertex may be reached from any other efficient vertex, such as the one found by the linear program above, by passing along some chain of neighboring efficient vertices (neighboring vertices are vertices connected by an edge). Then the way to find all efficient vertices is clear; simply check for efficiency the neighbors of each efficient vertex found. It happens that it is a simple matter to determine whether or not a vertex is efficient.

Having found all of the efficient vertices of the feasible region, the next step is to use these to express the efficient manifold. It has been noted that each efficient polyhedron in the efficient manifold can be represented by a set of efficient vertices. Then the efficient manifold can be represented by one or more sets of efficient vertices, where each set represents all convex combinations of the members of the set. Not all sets of efficient vertices specify efficient polyhedra, but a theorem attributed by Philip[1] to Kuhn and Tucker provides the ability to recognize whether or not a given set does. This theorem also provides the means for

ascertaining the efficiency of a given vertex. One procedure for searching through the possible combinations of efficient vertices is presented in Chapter V, and several more are presented in Chapter VI.

One rather troublesome difficulty which may arise is that the efficient manifold and the feasible region may turn out to be identical. It will be shown that a sufficient condition for this to be the case is for the set of criteria to be such that some positive weighting of them is identically zero. In this case, any arbitrary solution in the feasible region must be efficient because motion away in any direction is "down hill" with respect to at least one criterion, or if a direction is found which is not down hill with respect to any, then it is not up hill with respect to any either. This condition is easily checked.

Naturally, the efficient manifold is sensitive to the skill with which the model is built. So long as the criteria and constraints are correct and complete, the true optima will be in the efficient manifold somewhere. As in all mathematical programming, a poorly constructed constraint set may exclude the true optima, or it may include too much and cause the decision maker to settle upon an unattainable solution. If a criterion is omitted, the efficient manifold may be smaller than would otherwise be the case, and the true optima could be omitted. If too many criteria are employed, the efficient manifold may turn out to be unnecessarily large, and in unusual circumstances it may also eliminate portions of the efficient manifold arising from the problem with fewer criteria. In

[illegible]

[illegible]

[illegible]

[illegible]

this latter event, the new efficient manifold will still contain efficient solutions with scores identical to the omitted ones.

FOOTNOTES

[1]Philip, J. "Algorithms for the Vector Maximization Problem", *Mathematical Programming*, III (1972), pp. 207-229.

CHAPTER IV

THE SCHOOL OF HEALTH SERVICES EQUILIBRIUM PLANNING MODEL

IV-1 GENERAL

The previous chapter outlined an assortment of linear criteria and constraints which might prove useful in modeling enrollment and staffing decisions for an educational institution. In this chapter, some of these are used to build a model of one particular institution, the newly established School of Health Services of the Johns Hopkins University. Some of the history and characteristics of this interesting educational venture were summarized in Chapter II. The model herein was prepared primarily in order to provide a test of the analysis technique suggested in this thesis. It must be understood that although the author received considerable assistance from the Dean and members of his staff, neither the model nor its data bear any University endorsement, official or otherwise.

Several conceptual models of the School of Health Services were formulated before the following equilibrium model was decided upon. This model has several prominent advantages. It is mathematically much smaller and less complex than a dynamic model, and being a steady state model, it can not suffer transient effects. Perhaps most importantly, its outputs are much briefer and easier to

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

comprehend than are those which a dynamic model would yield. On the negative side, steady state is a prospect but dimly perceived by the staff of the School; growth is the focus of attention. Nevertheless, a steady state projection has some significance. It can expose the limits to growth which are posed by the resources available to the School and reveal which resource limitations will indeed operate to deter further growth. Knowing this, the Dean would then be enabled to direct his attention to ameliorating the potentially confining circumstances in advance. Financial projections for these steady states will also be illuminating. Finally, the disciplined and systematic enquiry which goes into developing such a model is valuable on its own account.

Because this model is a steady state model, some simplifications are possible over the dynamic case, which is implicit in some of the expressions and equations of the previous chapter. A steady state nature is imparted by presuming that faculty do not age or retire, and that student matriculations are just sufficient to balance graduations and dropouts, so that the total enrollments, and the enrollments in each year of each curriculum, are the same from year to year. The latter presumption permits the representative student's entire career to be telescoped into a single entity. Thus the model to follow counts only beginning students in each curriculum; each such student is credited with the tuition income he will generate and the resources he will require over the course of his passage through the School. The model is further simplified in that

student related variables and constants need only be differentiated according to curriculum, rather than by curriculum, by year within the curriculum, and by calendar year.

The model will be presented in three segments. General remarks and assumptions will come first; then criteria will be presented. Thirdly, the constraints are discussed. Data will be presented where it first arises incident to the criteria and constraints. After the discussion of the constraints is concluded, the mathematical formulae of the model will be summarized.

IV-2 ASSUMPTIONS AND NOTATION

On the advice of the staff of the School, the scope of the School of Health Services equilibrium planning model is limited to consideration of enrollments and staffing for the Health Assistant program and the impending baccalaureate programs for Health Associates, Nurse Practitioners, students of Health Services Management, and students of Environmental Hygiene. Each of these curricula is or will be two years long.

One very expensive class of faculty does not appear anywhere in the model to follow. These are the "faculty specialists"--the surgeons, microbiologists, pharmacologists, etc. They are excluded because in the judgment of the staff of the School their numbers will not vary significantly with changes in enrollments. Because of their high price and because instruction is not their only function (another primary responsibility is curricula guidance) they will be staffed in sufficient variety to represent the relevant specialties and in sufficient strength to present, without duplication, the relevant material in their specialties. Every effort will be made to adjust to changing enrollment patterns by changing the sizes of the groups which they address.

It was noted in Chapter II that primary teaching teams will be responsible for groups of 25 students each. There is one exception--the Health Assistant program. In this program, each primary teaching team will be responsible for a thirty student "section".

On consultation with the staff of the School, it was decided

to presume that no students will drop out, either temporarily or permanently. It is believed that careful selection of applicants and close counseling of enrollees will accomplish this. It is also presumed, on the strength of the diversity and relatively short duration of these programs, that--with one exception--there will be no student transfers among programs. The exception arises in the case of the Health Associate and Nurse Practitioner programs. These two programs differ principally in their emphases--the one being primarily concerned with ambulatory care, the other with institutional care. Current data estimates for the two programs are identical because the staff believes that the differences in emphasis will not be reflected in costs and facilities usages. Consequently these programs would be indistinguishable in the model, so their matriculations variables have been merged. According to the School's staff, a merger in fact is a distinct possibility.

The variables used in the School of Health Services Equilibrium Planning Model are as follows:

C_1 = Annual matriculations in the Health Assistant program

C_2 = Annual matriculations in the Health Associate and Nurse Practitioner programs

C_3 = Annual matriculations in the Health Services Management program

C_4 = Annual matriculations in the Environmental Hygiene program

Y_1 = The number of offices converted to dry laboratories

Y_2 = The number of hours weekly during which large classrooms are utilized in a partitioned mode (that is, used with the room divider closed, for two different classes).

Y_3 = The number of hours weekly during which large classrooms are filled by (single) medium or small classes.

Y_4 = The number of hours weekly during which medium classrooms are filled by small classes.

The symbols a and A, appropriately subscripted, will represent constants on the left and right sides of the constraints. The first subscript beneath each is the number of the constraint; that is, all a's and A's with the same first subscript belong to the same constraint. Where a second subscript occurs, its significance will be apparent from the accompanying text. The constants T_i represent income less expenses associated with matriculated students; the derivation of the T_i from familiar data is explained in section IV-3, in conjunction with the annual budget surplus criterion.

No faculty variables are required because no faculty categories are anywhere substitutable for one another. Consequently, given enrollments may be translated into exact faculty requirements, and vice versa. The known relations between enrollments and faculty are used to attribute faculty related requirements directly to the student variables.

IV-3 CRITERIA

Two of the sample criteria listed in section III-2 were selected for use in this model; they are numbers of graduates of each curriculum, and annual budget surplus. The remaining candidates are inappropriate for a variety of reasons. Faculty terminations can be projected only from dynamic models; this one is static. The unique modular method of instruction contains within it decisions on student/faculty ratios and on faculty discipline distributions, so these, also, are inappropriate criteria. Lastly, the staff of the School feels that there is little near term prospect of a building program expressly for the School; the capital investment criterion is thus superfluous.

Numbers of graduates:

All other things being equal, more graduates of a given curriculum are preferred to fewer. Since it is assumed that there will be no dropouts, the steady state number of graduates per year is equal to the number of students matriculated per year. Thus the criteria are:

Maximize Health Assistant program graduates = Max C_1

Maximize graduates of the Health Associate and Nurse Practitioner programs = Max C_2

Maximize graduates of the Health Services Management program = Max C_3

[illegible]

[illegible] [illegible]

[illegible]

[illegible]

Maximize graduates of the Environmental Hygiene program $= \text{Max } C_4$

Annual budget surplus:

Because net financial position is one of the major concerns associated with the birth of a new enterprise, the criterion "maximize instructional revenues less instructional expenses" is chosen as one of the criteria for this model. The most desirable financial measure would be the net annual financial position of the School, but, as noted in Chapter III, the sum of tuition and other student related incomes less the "proportional" costs (those which vary in proportion to numbers of enrollees) of instruction will suffice. The natural expression for this quantity would be the sum of the individual contributions of each student enrolled and each faculty member present during a given steady state year. But enrollments and numbers of faculty are not variables of this model; only matriculations are readily available. However, total enrollments and faculty size are directly related to matriculations, so the figure desired can be had by attributing the costs to the enrollments and faculty as before, and then attributing these to matriculations. Since it is assumed that there will be no dropouts, corresponding to each newly matriculated student will be one enrolled second year student. An equivalent approach would be to calculate the costs and benefits for each student throughout his passage through his curriculum; in this approach, the first and second years

[illegible]

[illegible] [illegible]

[illegible]

[illegible]

[illegible]

are thus represented just as they would be if the one year contributions of the first and second year students simultaneously present in a given year were computed. This device of attributing to the matriculating student his entire curricular requirements will be employed several times in this chapter. In this instance, it will of course include those figures pertaining to faculty whose presence is contingent on the student's presence somewhere in his curriculum. The criterion is:

$$\text{Maximize} \sum_{i=1}^{4} T_i C_i$$

where T_i is the income less expenses associated with each student matriculated in curriculum i. These are estimated to be:

$$T_1 = \$0$$

$$T_2 = \$-11,962$$

$$T_3 = \$-4644$$

$$T_4 = \$-4734$$

The first is presumed to be zero because it is expected that the National Institutes of Health will continue to fully fund the Health Assistant program. The other coefficients contain tuition and variable costs which are directly attributable to the student, plus each student's share of the variable costs attributable to faculty. Tuition for two years and a summer comes to $6750; allowing $200 a year

and $50 a summer for chemicals and laboratory specimens, computer budgets, and the like, leaves $6300. This is the student related portion of T_2, T_3, and T_4.

Primary teaching teams are presumed to be comparable in cost from one program to the next, though their professional composition will differ somewhat. The representative team for the Health Associate program includes a physician, at $27,000 annually, two social generalists at $15,000 annually each, and three health practitioners, at $13,300 each annually. Each such primary teaching team is supported by 1.3 secretaries, at $6800 per secretary per year. This comes to $105,740 in direct salary expenses. Adding 14%, the current Johns Hopkins fringe benefit figure, plus $200 per secretary per month for office supplies, plus $12,000 per team for media resource support, yields $135,664, the total variable cost per teaching team. Then, over the two year curriculum, the Health Assistant student is responsible for 2/30 of the annual costs of one such team, or $9044, and each student of the other programs is responsible for 2/25 of these costs, or $10,853.

Faculty specialists do not contribute to the T_1 because--as noted earlier--in the judgment of the staff of the School, their numbers will not vary significantly.

The final contribution to the T_1 is the per student costs of additional clinical supervisors. The numbers of days weekly that students in each year of each curriculum will receive clinical instruction is shown in the following table:

	H. Ass't program	H. Ao/N.P. programs	H. Serv. Mgmt. program	Env. Hygiene program
1st year students	1	1.5	1 hour	2 hours
2nd year students	3.25	2.83	1 hour	2 hours
Sum	4.25	4.33	.0533	.1066

Clinical supervision will be carried out in groups containing one instructor and two students, and each primary teaching team will perform the needed supervision of 8 of its own students (two for each of the four members qualified to perform clinical supervision). Then the average requirement (in days weekly) for additional clinical supervisors is 22/30 of the Health Assistant program figure above, or 17/25 of the figures for the other programs:

Health Assistant program:	3.117
Health Associate and Nurse Practitioner programs:	2.94
Health Services Management programs:	.0362
Environmental Hygiene programs:	.0725

Each clinical supervisor instructs two students at once, 5 days weekly; then his capacity is 10 student-days per week. Dividing the

previous row by 10 gives the following numbers of additional clinical supervisors required per student matriculated into the indicated curriculum:

Health Assistant programs:	.3117
Health Associate and Nurse Practitioner programs:	.294
Health Services Management program:	.0036
Environmental Hygiene program:	.0072

Full time equivalent clinical supervisors will be either physicians, salaried at roughly $27,000 annually, or health practitioners at about $13,300 annually. On the assumption that their relative proportions will be roughly 18 of the former to 10 of the latter, the average direct salary becomes $22,107; adding 14% for fringe benefits yields a net cost of $25,202 per additional clinical supervisor. It is presumed that the office and secretarial support requirements of additional clinical supervisors are negligible. Multiplying the figures in the previous line by this sum gives:

Health Assistant program:	$7855
Health Associate and Nurse Practitioner programs:	$7409

Health Services Management

program: $91

Environmental Hygiene

program: $181

Summarizing, the income less proportional costs of instruction per matriculated student are shown in the following table:

	Health Assistant program	H. Associate/ N. Practitioner programs	H. Serv. Mgmt. program	Env. Hygiene program
Direct student revenues less direct student costs	$16,899*	$6300	$6300	$6300
Primary teaching team costs	$-9044	$-10853	$-10853	$-10853
Additional clinical supervisor costs	$-7855	$-7409	$-91	$-181
Total	$0	$-11,962	$-4644	$-4734

*Includes $10,599 from the NIH.

[illegible]

[illegible]

[illegible]

[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

[illegible]

IV-4 CONSTRAINTS

Two of the constraint families listed and discussed in section III-3 are encorporated into the School of Health Services equilibrium planning model; they are faculty availability and facilities availability. The other listed constraint types are inappropriate either because they are obviated by the construction of the model (e.g. the mathematical identities relating faculty and students, or matriculations and enrollments), because there is no foreseeable prospect of violating the availability of the resource in question (student applications), or because the constraint is simply not an issue for the School (faculty or student body composition).

The only faculty category which in the long run is anticipated to be in short supply is part time clinical supervisors. These will be called "additional clinical supervisors", for their supervision is in addition to that provided by members of the primary teaching teams. Thus the model includes one faculty availability constraint. The facilities availability family, on the other hand, contains eight constraints because eight different facility types which might conceivably be in short supply have been identified. These are faculty offices, three types of laboratories, and four types of classrooms.

Availability of additional clinical supervisors:

It appears that the only faculty category whose availability

might possibly be limiting is additional clinical supervisors. The appropriate constraint is

$$\sum_{i=1}^{4} a_{1i}C_i \leq A_1$$

where a_{1i} is the number of additional clinical supervisors required per student matriculated in the i^{th} curriculum and A_1 is the number of additional clinical supervisors realistically available to the School, in full time equivalents. The estimates for the a_{1i} are:

$$a_{11} = .4033$$

$$a_{12} = .408$$

$$a_{13} = .0362$$

$$a_{14} = .0725$$

These were developed as follows. It is presumed that clinical supervision is carried out in groups of one clinical supervisor and two students. Of the six member primary teaching team, four members will provide clinical supervision for students of their sections; since all members of a section undergo clinical supervision at the same times, this means that the primary teaching team can accommodate the clinical supervisory requirements of eight students, leaving 22 students from each health assistant section or 17 students from each section in other curricula who will need to be accommodated by additional clinical supervisors. Clinical supervision is most intense

[illegible]

$$[illegible]$$

[illegible]

[illegible]

[illegible]

during the spring terms; thus it is on the requirements at these times that the estimates are based. The expected spring term intensity of clinical supervision in days weekly for students of each year of each curriculum is shown in the following table.

	H. Ass't.	H. Ao/N.P.	H. Serv. Mgmt.	Env. Hygiene
1st spring term	1	1.5	1 hour	2 hours
2nd spring term	4.5	4.5	1 hour	2 hours
Sum	5.5	6	.0533	.1066

Multiplying by 22/30 or 17/25 as appropriate yields the average number of days weekly each such student requires an additional clinical supervisor.

H. Ass't.	H. Ao/N.P.	H. Serv. Mgmt.	Env. Hygiene
4.033	4.08	.0362	.0725

Each full time equivalent additional clinical supervisor instructs two students at once, five days weekly; his capacity is 10 student days weekly. Then dividing the previous row by 10 gives the number of additional clinical supervisors required per graduate of each program; these are the figures cited earlier.

A_1, the number of additional clinical supervisors available, in full time equivalents, is estimated to be 28. This is based on the presumptions that each interested individual contributes 1/10

full time equivalent--i.e. a morning or an afternoon a week, that 25% of the Johns Hopkins house staff in Pediatrics, Medicine, Surgery, Psychiatry, and Obstetrics (totalling 399) will be interested, and that 10% of the general practitioners (420), internists (654), pediatricians (399), and OB-GYN specialists (365) in the Baltimore metropolitan area will be interested. The hospital staff figures are as of September 1972; the Baltimore metropolitan area figures are from an American Medical Association survey report for the end of 1969, the last year in which the AMA cumulated and published its survey by metropolitan areas. Apparently hospital house staff are not included in the survey.

Availability of faculty offices:

The appropriate constraint is

$$\sum_{i=1}^{4} a_{2i}C_i + 2Y_1 \leq A_2$$

where a_{2i} is the number of faculty desks required per student matriculated in curriculum i and where A_2 is the total number of desks available. The possibility exists that offices might need to be converted to dry (instrumentation) laboratories; the term $2Y_1$ reflects this drain upon the number of available desks, assuming there are two desks per office. It is presumed that only the primary teaching teams and additional clinical supervisors are proportional to enrollments, and that of these, only primary teaching team

members will require desks. In the Health Assistant program, each student's share of primary teaching team members is 1/30 of 6 for each year, or .4 for the two years; in the other curricula, it is 1/25 of 6 for each of two years, or .48. Then

$$a_{21} = .4$$

$$a_{22} = a_{23} = a_{24} = .48$$

Hampton House, formerly a student nurse's residence, is currently undergoing a prolonged renovation; when complete, this building will house the School of Health Services and portions of other divisions of the University. Renovation blueprints for the floors to be occupied by the School show space for 80 desks for teaching faculty. Approximately ten of these will be required for selected faculty specialists; then $A_2 = 70$.

Facilities usage information:

Six types of instructional facilities (three sizes of classroom, and three types of laboratory) have been identified as potentially being in short supply. The classroom types are large classrooms--that is, classrooms capable of holding a double section, consisting of 50 to 60 students, depending on the curriculum; medium classrooms--i.e., classrooms capable of accommodating a single section but not two; and small classrooms--i.e., those capable of holding a half section but not an entire section. Bench laboratories

are "wet" laboratories fitted with laboratory stations for each person; these labs are of the sort used for practical instruction in chemistry, biology, etc. It is anticipated that entire student sections will perform their bench laboratory studies at once; thus the capacity of a bench laboratory must be 25 to 30 students. Demonstration laboratories are wet laboratories in which the student will observe laboratory experiments and procedures performed by an instructor. Demonstration laboratory sessions will be attended by half sections; therefore, the capacity of each demonstration laboratory should be at least 13 to 15 students. Dry laboratories are rooms furnished with special equipment required in practical instruction, for example a bed on which one student may recline while another takes his pulse and blood pressure. It is anticipated that dry laboratories will be utilized by groups of four students.

The following table contains the numbers of hours weekly that students of each curriculum are expected to spend under instruction in the indicated type of facility. The figures are for the fall semester of the year because usage of these facilities is heavier during the fall than during the spring. The symbol I denotes the fall of the first year of the indicated curriculum; II denotes the fall of the second year.

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

		Classrooms:			Laboratories:		
		Large	Medium	Small	Dry	Demo.	Bench
H. Ass't.	I	2	5	5	1	2	1
	II	2	5	4	1	2	1
H. Assoc/ N. Prac.	I	7	3	5	3.75	1.87	1.87
	II	7	3	5	0	1	1
H. Serv. Mgmt.	I	7.5	22.5	3.75	3.75	0	0
	II	7.5	22.5	3.75	3.75	0	0
Environ. Hygiene	I	7.5	15	11.25	1.87	1.87	3.75
	II	1.87	3.75	3.75	0	0	0

FACILITIES USAGE TABLE

(hours weekly)

<u>Dry (instrumentation) laboratory availability:</u>

The appropriate constraint is:

$$\sum_{i=1}^{4} a_{3i}C_i - 37.5Y_1 \leq A_3$$

where a_{3i} represents the number of hours (weekly) of dry laboratory availability required per student matriculated in the i^{th} curriculum, and A_3 represents the number of hours of dry laboratory usage available weekly. The term $-37.5Y_1$, which is based on a 7.5 hour school day, represents the contribution in hours per week of each room which is nominally a faculty office but which is converted to dry laboratory use.

Dry laboratory usage rates per student are the figures from the facilities usage table summed over the length of each curriculum and divided by four (reflecting the expectation that dry laboratories will be used by four students at once). Thus

$$a_{31} = .5$$

$$a_{32} = .94$$

$$a_{33} = 1.875$$

$$a_{44} = .468$$

The availability of each type of laboratory is not presently known. No laboratories of any kind are planned for Hampton House; excess capacity elsewhere on the Johns Hopkins East Baltimore campus

is difficult to assess. It is not generally possible to estimate excess laboratory capacity by tallying up unused rooms, because excess capacity usually manifests itself in a spreading out of ongoing activities and in creation of new ones which would not have been undertaken at all had "free space" not been available. Nevertheless, it is known that some facilities will be made available to the School, and certainly it can be assumed that the closure of the Hospital's nursing program will result in some free space. A reasonable minimum estimate is one laboratory of each type continuously available.

$$\text{Then } A_3 = 37.5.$$

Demonstration laboratory availability:

The appropriate constraint is:

$$\sum_{i=1}^{4} a_{4i} C_i \leq A_4$$

where a_{4i} is the number of hours weekly of demonstration laboratory usage availability required per student matriculated in curriculum i, and where A_4 is the total availability of demonstration laboratories, measured in hours weekly.

Demonstration laboratory requirements per student are the figure from the facilities usage table summed over the length of the curricula and divided by 12.5 or 15, reflecting the expectatation that demonstration laboratories will be used by half sections. Then

$$a_{41} = .266$$

$$a_{42} = .227$$

$$a_{43} = 0$$

$$a_{44} = .150$$

The difficulty in estimating available laboratory space has been mentioned in conjunction with the previous constraint. It is felt that at least one full time demonstration laboratory or its equivalent in part time availability will be available; then $A_4 = 37.5$.

Bench laboratory availability:

The appropriate constraint is

$$\sum_{i=1}^{4} a_{5i}C_i \leqq A_5$$

where a_{5i} is the amount of bench laboratory usage required (in hours per week) per graduate of curriculum i, and where A_5 is the expected availability of bench laboratories, in hours per week.

Bench laboratory requirements per student are found by summing the figures in the facilities usage table over the length of each curriculum and dividing by 25 or 30 as appropriate (recall that bench laboratories will be occupied by entire student sections). Then

$a_{51} = .0667$

$a_{52} = .115$

$a_{53} = 0$

$a_{54} = .15$

Again, in the absence of reliable information on the ultimate availability of laboratory space, one full time laboratory is presumed; $A_5 = 37.5$.

Large classroom availability:

The appropriate constraint is

$$\sum_{i=1}^{4} a_{6i}C_i+Y_2+Y_3 \leqq A_6$$

where a_{6i} is the requirement in hours per week for large classroom space per student entering the i^{th} curriculum, and where A_6 is the availability of large classrooms in hours per week. The terms $+Y_2$ and $+Y_3$ reflect the effects of using large classrooms for classes other than (combined) double section classes (recall that Y_2 is the number of hours weekly that large classrooms are partitioned and used for smaller classes, and Y_3 is the number of hours weekly that large classrooms are used for single smaller classes).

Large classroom requirements are found by summing the figures from the facilities usage table over the length of each curriculum and dividing by 50 or 60 as appropriate; then

$a_{61} = .0667$

$a_{62} = .28$

$a_{63} = .30$

$a_{64} = .187$

Hampton House renovation blueprints show one large seminar of approximately 630 net square feet (this room is to be equipped with a divider) and three of 560 net square feet each; any of these four classrooms could accommodate a double section (though perhaps not very comfortably). Then $A_6 = 4 \times 37.5 = 150$ hours weekly.

Medium classroom availability:

The appropriate constraint is

$$\sum_{i=1}^{4} a_{7i}C_i - 2Y_2 - Y_3 + Y_4 \leqq A_7$$

where a_{7i} is the requirement in hours per week for medium classroom space per student entering the i^{th} curriculum, and where A_7 is the availability of medium classrooms in hours per week. The term $-2Y_2$ reflects the effect of using a partitioned large classroom for two medium classes simultaneously; the term $- Y_3$ reflects using a large classroom for a single medium class. The last term, $+Y_4$, incorporates the utilization of medium classrooms for small (half section) classes.

Medium classroom requirements per student are found by

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

summing the figures in the facilities usage table over the curricula and dividing by 30 or 25 as appropriate. Then

$$a_{71} = .333$$

$$a_{72} = .24$$

$$a_{73} = 1.8$$

$$a_{74} = .75$$

Hampton House renovation blueprints show one seminar of approximately 400 net square feet, two of 390 net square feet, another of 360, one of 330, and two of 315 net square feet each. Each of these seven rooms is capable of accommodating a full section; then $A_7 = 7 \times 37.5 = 262.5$ hours weekly.

Small classroom availability:

The appropriate constraint is

$$\sum_{i=1}^{4} a_{8i}C_i - Y_4 \leqq A_8$$

where a_{8i} is the requirement in hours per week of small classroom availability per student entering the i^{th} curriculum, and A_8 is the availability of small classrooms in hours per week. The term $-Y_4$ reflects the relaxation of this constraint which occurs when Y_4 weekly hours of medium classroom availability is devoted to small classes.

Small classroom requirements per student matriculated in program i are found by summing the figures in the facilities usage table over the length of each curriculum and dividing by 12.5 or 15 as appropriate. Then

$$a_{81} = .6$$

$$a_{82} = .8$$

$$a_{83} = .6$$

$$a_{84} = 1.2$$

Hampton House renovation blueprints show one seminar of approximately 210 net square feet; this room will accommodate one half student section. Then $A_8 = 37.5$.

Availability of partitionable large classrooms:

During the discussion of the large classrooms availability constraint it was noted that one of the large classrooms is to be equipped with a divider; then the amount of partitioned large classroom usage, represented by the variable Y_2, can not exceed the number of hours available weekly in this one partitionable room. Then the appropriate constraint is

$$Y_2 \leq 37.5$$

[illegible]

$$[illegible]$$

[illegible]

[illegible]

IV-5 THE MATHEMATICAL MODEL

The following is a mathematical summary of the School of Health Services equilibrium planning model, presented, for convenience, in matrix format:

$$\text{Maximize} \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & -11962 & -4644 & -4734 \end{bmatrix} \begin{bmatrix} C_1 \\ C_2 \\ C_3 \\ C_4 \end{bmatrix}$$

Subject to

$$\begin{bmatrix} .4033 & .408 & .0362 & .0725 & 0 & 0 & 0 & 0 \\ .4 & .48 & .48 & .48 & 2 & 0 & 0 & 0 \\ .5 & .94 & 1.875 & .468 & -37.5 & 0 & 0 & 0 \\ .226 & .227 & 0 & .150 & 0 & 0 & 0 & 0 \\ .0667 & .115 & 0 & .15 & 0 & 0 & 0 & 0 \\ .0667 & .28 & .30 & .187 & 0 & 1 & 1 & 0 \\ .3333 & .24 & 1.8 & .75 & 0 & -2 & -1 & 1 \\ .6 & .8 & .6 & 1.2 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} C_1 \\ C_2 \\ C_3 \\ C_4 \\ Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \end{bmatrix} \leqq \begin{bmatrix} 28 \\ 80 \\ 37.5 \\ 37.5 \\ 37.5 \\ 150 \\ 262.5 \\ 37.5 \\ 37.5 \end{bmatrix}$$

The results of the actual manipulation of this model will be presented in Chapter VII following a detailed presentation, in Chapters V and VI, of the general method of finding the efficient manifold.

[illegible]

[illegible]

[illegible]

[illegible]

CHAPTER V

METHOD OF SOLUTION

V-1 INTRODUCTION

This chapter is concerned with finding solutions to the following general problem, which will be called "problem P":

Find the set of efficient solutions to

$$(\text{Max}) \quad \sum_{j=1}^{n} c_{1j}x_j$$

$$\sum_{j=1}^{n} c_{2j}x_j$$

$$\vdots$$

$$\sum_{j=1}^{n} c_{kj}x_j$$

subject to:

$$\sum_{j=1}^{n} a_{1j}x_j = b_1$$

$$\sum_{j=1}^{n} a_{2j}x_j = b_2$$

$$\vdots$$

$$\sum_{j=1}^{n} a_{pj}x_j = b_p$$

$$\sum_{j=1}^{n} a_{p+1,j}x_j \leqq b_{p+1}$$

$$\vdots$$

$$\sum_{j=1}^{n} a_{mj}x_j \leqq b_m$$

$$-x_1 \leqq 0$$

$$-x_2 \leqq 0$$

$$\vdots$$

$$-x_n \leqq 0$$

where it is assumed that the set of all feasible solutions to the constraint set is bounded, and that there do not exist positive scalars L_i such that

$$\sum_{i=1}^{k} L_i c_{ij} = 0 \quad \text{for} \quad j = 1,2,\ldots n$$

It was noted in Chapter III that efficient solutions to this problem are of interest because the preferred solution must be among them. In this chapter, a technique for finding the complete set of efficient solutions to problem P will be presented. The reader with no taste for algorithms or mathematical proofs is invited to skip to Chapter VII where he will find an account of the application of this technique to the School for Health Services equilibrium planning model which was developed in Chapter IV.

The procedure for finding the efficient points will be presented in two parts. First, an algorithm for finding all efficient vertices of the polyhedron of feasible solutions will be presented. It will be demonstrated that the set of all efficient solutions can be expressed in terms of these efficient vertices. Then the second part of the procedure, an algorithm for finding such an expression, will be presented.

[illegible]

[illegible]

V-2 BASIC DEFINITIONS AND NOTATION

The following notation and basic definitions will be employed in the remainder of this chapter and in Chapter VI.

Inequality notation: Let $x = (x_1, x_2, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ be two vectors in n-space.

(1) $x \geqq y$ if $x_i \geqq y_i$ for all $i = 1,2,\ldots,n$.

(2) $x \geq y$ if $x \geqq y$ and $x_i > y_i$ for at least one i.

(3) $x > y$ if $x_i > y_i$ for all $i = 1,2,\ldots,n$.

Transpose notation: When t is employed as a superscript it will indicate the transpose operation on the superscripted entity.

Inner product notation: Without ambiguity, xy will signify simple multiplication - i.e. "x times y" when x or y or both are scalars, and will signify the inner product of x and y when both are vectors or matrices.

Components of vectors: Components of vectors (and rows or columns of matrices, when specified) will customarily be represented by subscripting the symbol for the vector (matrix); e.g., x_i is the i^{th} component of vector x. Occasionally, components of the results of inner product operations will be referred to in the same fashion. For example, if L is a vector and G is a matrix of appropriate dimension then the i^{th} component of the vector resulting from the inner product of L and G may be written $(LG)_i$.

"1" and "0": It is sometimes necessary to employ a vector every component of which is one, or a vector every component of which is zero. The symbols 1 and 0 will be used, respectively. In every case the information in the remainder of the equation or expression will be sufficient to infer whether these vectors or the scalars zero and one are intended, and in the former case the length of the vectors will be clear also.

Problem P (vector and matrix notation):

Define: x = the column vector of variables x_i. x has length n.

L = the row vector (of length k) of multipliers L_i.

C = the matrix of criteria coefficients c_{ij}. C has k rows and n columns. Notice that the coefficients of each particular criterion appear as a row of C. It will be assumed that there does not exist $L > 0$ such that $LC = 0$.

F = the "feasible region" - that is, the set of all x satisfying the given constraints:

$$\sum_{j=1}^{n} a_{ij}x_j = b_i \qquad i = 1,\ldots,p$$

$$\sum_{j=1}^{n} a_{ij}x_j \leq b_i \qquad i = p+1\ldots m$$

$$-x_i \leq 0 \qquad i = 1\ldots n$$

It is assumed that F is bounded.

A = the matrix of coefficients on the left of the ='s in the equation set above. A has dimensions $(m+n)$ by n.

b = the column vector (of length $m+n$) on the right of the ='s above.

In this notation P is simply:

Find the efficient points to

(Max) Cx

subject to x in F

Efficiency: Feasible solution x will be said to be efficient to problem P if and only if there does not exist y in F such that

$$y \geqq x$$

Vertex: The terms "vertex", "extreme point", and "corner" will be used interchangeably to denote the solution to a "regular system of tight constraints" (see any linear programming text).

Remark: It is well known that any point in a bounded polyhedron (such as F) may be expressed as a convex combination of its vertices. For example, if x is in F and if $x^1, x^2, \ldots, x^r$ are the vertices of F, then there exist multipliers n_i such that

$$x = \sum_{i=1}^{r} n_i x^i$$

$$\sum_{i=1}^{r} n_i = 1$$

$$n_i \geqq 0 \qquad i = 1,2,\ldots r$$

Also, the intersection of a polyhedron and of one of its constraints taken as an equality constraint is another polyhedron, whose vertices are vertices of the original polyhedron. (The geometry of polyhedra is well explained in Simmonard.[1])

Tight constraint: A constraint (whether originally stated as an equality or as an inequality) which is satisfied as an equality is said to be "tight".

Facet: A facet is the intersection of a polyhedron and one or more of its tight constraints.

Remark: It may be that a facet is vacuous - if, for example, the chosen tight constraints do not intersect in F. By the remarks above, when a facet is nonvacuous, it is a polyhedron every point of which may be expressed as a convex combination of its vertices - that is, of some of the vertices of F. Clearly every nonvacuous facet arising from F must contain at least

one vertex. The symbol $f(J)$ will denote the facet consisting of the intersection of F and the tight constraints whose indices are in the set J. Notice that if set J contains set J' then facet $f(J')$ contains facet $f(J)$, for $f(J)$ is the intersection of $f(J')$ and those tight constraints indexed in J but not in J'. Incidently, vertices are facets, for they are intersections of F and of a regular system of tight constraints.

Efficient facet: An efficient facet is a facet of which every solution x in the facet is efficient.

Efficient manifold: The efficient manifold is the set of all efficient solutions to problem P.

V-3 PRELIMINARY THEOREMS

The literature of multicriterion optimization was reviewed in Chapter I: the fundamental results on which the efficient point search algorithms are based are presented in this section.

Problem P includes the assumption that there does not exist $L > 0$ such that $LC = 0$. The following theorem establishes that the problem is uninteresting when this condition is not met.

Theorem V-1*: Every x in F of problem P is efficient if there exists $L > 0$ such that $LC = 0$.

Proof: By Steimke's theorem of the alternative[2] the existence of row vector $L > 0$ such that $L = 0$ implies that there does not exist x such that $Cx \leq 0$. Since $LC = 0$, then $L(-C) = 0$, and there does not exist x such that $Cx \geq 0$. Consider any two solutions x' and x" such that $x' \neq x''$. Define $x = x''-x'$. It has been seen that $Cx \not\leq 0$ and $Cx \not\geq 0$. Notice that $Cx \not\leq 0$ implies either

a) $Cx = 0$, or

b) $(Cx)_i > 0$ for some i.

Also, $Cx \not\geq 0$ implies either

c) $Cx = 0$, or

d) $(Cx)_j < 0$ for some j.

---

*Throughout this chapter and the next, theorems, lemmas, and corollaries not expressly credited to someone else were derived by the author and are, to the best of his knowledge, original.

Together these imply either

e) $Cx = 0$, or

f) $(Cx)_i > 0$ for some i and $(Cx)_j < 0$ for some j.

That is (from $x = x''-x'$) either

g) $Cx'' = Cx'$, or

h) $(Cx'')_i > (Cx')_i$ for some i and $(Cx'')_j < (Cx')_j$ for some j.

Then $Cx'' \not\geq Cx'$. Since x" was arbitrary, x' must be efficient.

Since x' was arbitrary, all of F is efficient.

QED

It happens that the existence of $L > 0$ such that $LC = 0$ implies and is implied by the existence of $L \geqq 1$ such that $LC = 0$. Let vector $M = L - 1$ and note that the assumption may be reexpressed "there does not exist $M \geqq 0$ such that $MC = -LC$". This condition is in a form easily tested by phase one of the two phase simplex method.

The following two theorems date in principle back to 1951, to Kuhn and Tucker's fundamental paper on nonlinear programming. These, and corollary V-3-1, are cited and proven by Philip.[3] Problem P' is defined to be problem P modified to include only "$\leqq$" constraints.

Theorem V-2 (Philip): Let $w = LC$ for some $L > 0$. If x solves Max wx subject to x in F, then x is efficient to problem P'.

Theorem V-3 (Philip): Let J be the set of indices of the rows (a_i) of matrix A of problem P' which are tight at some solution x. This x is efficient to problem P' if and only if there exist row vector $L > 0$ and multipliers $k_i \geqq 0$ (for all i in J) such that

$$LC = \sum_{\text{all } i \text{ in } J} a_i k_i$$

Colollary V-3-1 (Philip): Under the conditions above, x is efficient to problem P' if and only if there exists $L \geqq E > 0$ for arbitrary vector E, and $k_i \geqq 0$ for all i in J such that

$$LC = \sum_{\text{all } i \text{ in } J} a_i k_i$$

The following corollary extends this necessary and sufficient condition to problem P.

Corollary V-3-2: Let J be the set of indices of tight constraints at x in F of problem P. Solution x is efficient to problem P if and only if there exist $L \geqq E > 0$ and k_i for all i in J, where $k_i \geqq 0$ if constraint i is an inequality and unrestricted if constraint i is an equality in the statement of F, such that

$$LC = \sum_{\text{all } i \text{ in } J} a_i k_i$$

Proof: (If clause) The equations defining F may be grouped into an equality set and an inequality set; letting A^1, b^1, A^2, and b^2 be the coefficient and right hand side matrices and vectors of these respectively, F may be expressed:

$$A^1x = b^1$$
$$A^2x \leqq b^2$$

An equivalent expression for F is:

$$A^1x \leqq b^1$$
$$A^2x \leqq b^2$$
$$-A^1x \leqq -b^1$$

With F specified in this all inequality manner, the problem is ready for application of corollary V-3-1; that is, solution x is efficient if and only if there exists row vector $L \geqq E > 0$ and nonnegative multipliers k'_i for all i in J(and in the third set of equations above), such that

$$LC = \sum_{\text{all } i \text{ in } J} a_i k'_i$$

These k_i' may correspond to equations in any of the three groups of equations in the inequality expression for F. Let J be partitioned into sets J^1, J^2, and J^3 accordingly. Then the previous expression becomes:

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

$$LC = \sum_{i \text{ in } J^1} a_i k_i' + \sum_{i \text{ in } J^2} a_i k_i' + \sum_{i \text{ in } J^3} a_i k_i'$$

All equations in the first and third groups of equations in the inequality expression for F must necessarily be satisfied as equalities, so for every a_i in the first group, there is a term in the first summation of the last equation, and also another for $-a_i$ in the third summation. Grouping these two, making the sign substitution, and representing k_i' from J^3 by k_i'' gives:

$$LC = \sum_{i \text{ in } J^1} a_i(k_i' - k_i'') + \sum_{i \text{ in } J^2} a_i k_i'$$

Setting $k_i = (k_i' - k_i'')$ from the first group and $k_i = k_i'$ for i from J^2 gives a set of k^i satisfying

$$LC = \sum_{i \text{ in } J} a_i k_i$$

$k_i \geqq 0$ when constraint i is an inequality

This is the condition specified in the statement of the corollary.

(Only if clause) Since every real multiplier k_i can be expressed as the difference of two nonnegative multipliers k_i' and k_i'', a set of k_i satisfying the condition in the statement of the theorem can be used to create a set of k_i' satisfying the corresponding conditions for the all inequality version of problem P'.

QED

The following definitions and concepts are familiar from linear programming; the only novelty is the occasional addition of an extra subscript reflecting the fact that the cost vector is now a cost matrix. Let:

B = any basis (later on, bases will be identified by subscripts, e.g. B_i)

c_i = the i^{th} row of objectives matrix C

$c^B{}_i$ = the vector of those components of c_i corresponding to the basic components of x for the basis B.

$a_{\cdot j}$ = the j^{th} column of matrix A

$y_j = B^{-1}a_{\cdot j}$

Y = the matrix of columns y_j

$z_{ij} = c^B{}_i y_j$

$g_{ij} = c_{ij} - z_{ij}$ This is the reduced cost of the j^{th} component of x calculated with the current basis and the i^{th} row of C.

G = the matrix of reduced costs g_{ij}

$d = B^{-1}b$ This is the vector of values of the basic variables.

The following tableau, the usual simplex tableau with a reduced costs matrix instead of a reduced costs row, will be styled the "main tableau";

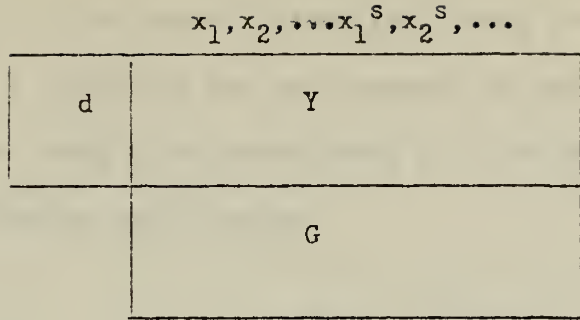


Figure III

Philip mentions without proof that the existence of $L > 0$ such that $LG \leqq 0$ is an efficiency criterion, analagous to the optimality criterion of linear programming. This condition will be referred to as the "efficiency criterion". Lemmas V-4 and V-5 and theorem V-6 establish the validity of this criterion, and expose an idiocyncracy unmentioned by Philip.

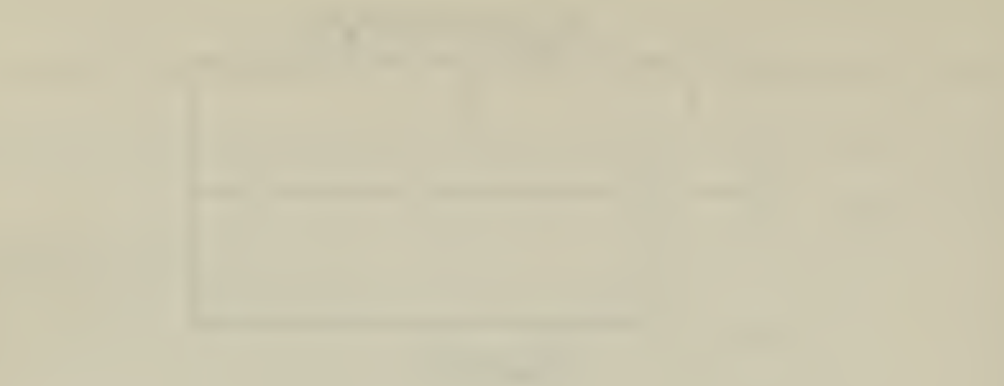
<u>Lemma V-4</u>: If there exists row vector $L > 0$ such that $LG \leqq 0$ for the matrix G corresponding to a tableau representing solution x to problem P, then x is efficient to problem P.

Proof: Follows directly from theorem V-2.

<u>Lemma V-5</u>: If x is efficient to problem P then there exists vector $L > 0$ such that x solves Max LCx subject to x in F.

Proof: By the corollary to Philip's third theorem, x is efficient implies there exists $L > 0$ and multipliers k_i ($\geqq 0$ when constraint i is an inequality in the statement of problem P) such that

$$LC = \sum_{i \text{ in } J} a_i k_i$$



where J is the set of constraints which are tight at x. The vector L satisfies the requirements of the lemma, for if it did not, then there would exist x' in F such that LCx' > LCx. Substituting the relation

$$LC = \sum_{i \text{ in } J} a_i k_i$$

into the above expression yields

$$\Big(\sum_{i \text{ in } J} a_i k_i\Big)x' > \Big(\sum_{i \text{ in } J} a_i k_i\Big)x = \sum_{i \text{ in } J} k_i(a_i x)$$

or, since $a_i x = b_i$ for i in J, since J is the set of tight constraints at x,

$$\sum_{i \text{ in } J} k_i(a_i x') - \sum_{i \text{ in } J} k_i(b_i) > 0$$

or

$$\sum_{i \text{ in } J} k_i(a_i x' - b_i) > 0$$

k_i may be negative only when constraint i is an equality; then $a_i x' - b_i = 0$, so no term containing $k_i < 0$ contributes to the sum above. Then there must be at least one term with $k_i > 0$ and $a_i x' - b_i > 0$ which is a contradiction since x' is feasible.

QED

Theorem V-6: If solution x is an efficient vertex to problem P, there exists a basis B corresponding to x such that for the

associated matrix G there exists vector $L > 0$ such that $LG \leqq 0$.

Proof: By the previous lemma, there exists vector $L > 0$ such that x solves max LCx subject to x in F. In a concluding tableau of this problem corresponding to x the reduced costs row is nonpositive. That is to say:

$$(LC)_j - (LC)^B y_j \leqq 0 \text{ for all } j$$

or

$$\sum_{i=1}^{m} c_{ij} L_i - (\sum_{i=1}^{m} L_i c_i{}^B) y_i \leqq 0 \text{ for all } j$$

or

$$\sum_{i=1}^{m} L_i (c_{ij} - c_i{}^B y_j) \leqq 0 \text{ for all } j$$

or, using the relation $g_{ij} = c_{ij} - c_i{}^B y_j$

$$\sum_{i=1}^{m} L_i g_{ij} \leqq 0 \text{ for all } j$$

or

$$LG \leqq 0$$

QED

These results are not the same as saying that x is efficient if and only if such L exists, for when the vertex at x is degenerate, there are several corresponding bases and corresponding reduced costs matrices. The last theorem proves only that one of them must satisfy the efficiency criterion; an example containing a degenerate vertex for one basis for which such an $L > 0$ does not exist is

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

$$[illegible]$$

$$[illegible]$$

$$[illegible]$$

[illegible]

$$[illegible]$$

[illegible]

[illegible]

[illegible]

given in Appendix I. Then the existence of such an L for a given G is sufficient but not necessary for the efficiency of the vertex represented.

V-4 AN ALGORITHM FOR GENERATING ALL EFFICIENT EXTREME POINTS

Theorem V-2 provides an easy method for finding an initial efficient vertex. The procedure is simply to take some arbitrary strictly positive weighting of the rows of C, using it as the objective gradient of an ordinary linear program (to be maximized over F). The simplex will terminate at a vertex, which must be efficient, according to the theorem.

Philip[3] presents a method, based on the simplex tableau, for determining whether a given efficient vertex is the only efficient vertex, and if not, for determining another vertex, one which may be reached by one pivot. (A vertex or basis which may be reached from another by one simplex pivot will be said to "neighbor" the first). In the algorithm to follow, this procedure will be used to find all neighboring efficient vertices of the initial vertex, and then all neighbors of those neighbors, and so on. It will be proven that all efficient vertices can be generated by this procedure.

Suppose that the linear program "Max LCx subject to x in F" is solved for some fixed positive vector L. At termination there will be a basis B for which the associated reduced costs matrix G has the property that there exists $L > 0$ such that $LG \leqq 0$. It is desired to know whether any neighboring vertices are efficient also. The following procedure for answering this question is due to Philip.

Suppose that there exists vector L' such that $L' > 0$, $L'G \leqq 0$, and $(L'G)_i = 0$ for some i. Then both the vertex represented

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

by the present tableau and that represented by the tableau which would result if column i were pivoted into the basis solve the problem "Max $L'Cx$ subject to x in F". (Clearly the current basis solves this problem, for $L'G \leqq 0$ implies the linear programming optimality criterion is met; when column i is pivoted in, since $(L'G)_i = 0$, the linear programming reduced costs row is unchanged; the new tableau meets the optimality criterion also).

It happens that the new tableau may be simply another expression for the same vertex. In this case, the vertex is degenerate, and it can happen that not all neighbors of a degenerate vertex are uncovered in the course of investigating all bases neighboring a given basis. Nevertheless, the procedure outlined will in the end find all efficient vertices, as theorem V-10 will show.

Thanks to the homogeneity of the requirements $L > 0$ and $LG \leqq 0$, if such L exists then there exists $L \geqq E > 0$ for arbitrary positive E, and vice versa. Then an equivalent efficiency condition is the existence of $L \geqq 1$ such that $LG \leqq 0$. Let $M_i = L_i - 1$ and perform a change of variables; then the condition becomes: the vertex represented is efficient if there exists $M \geqq 0$ such that $MG \leqq 1G$. This condition is easily checked through phase one of the two phase simplex. In checking the efficiency of a neighbor, it is desired to find such an M with the additional property that $(MG)_i = (1G)_i$ for the particular column i in question. This, also, is easily checked; simply minimize the value of the slack to the constraint $(MG)_i = (1G)_i$ in the problem just outlined; if the problem

is feasible and the slack can be driven to zero, then the desired condition is met; the basis resulting from the introduction of column i will satisfy the efficiency criterion.

Since it is desired to find all efficient neighbors, several different pivot locations may need to be investigated for any given column i satisfying the entry criterion just discussed. The exit criteria of interest are the following:

$$\text{(a)}\ d_j/y_{ji} = \underset{s/y_{si} > 0}{\text{minimum}}\ (d_s/y_{si})$$

$$\text{(b)}\ d_j = 0 \text{ and } y_{ji} \neq 0$$

The reader may verify that for column i satisfying the entry criterion just discussed, any pivot row satisfying either of the conditions (a) or (b) above will give a new tableau whose reduced costs row (for the objective L'Cx) will satisfy the linear programming optimality criterion--thus the associated G satisfies the efficiency criterion.

The following lemmas and theorem V-10 establish that a neighbor by neighbor search of the type suggested will generate all vertices efficient to problem P.

Lemma V-7 ("Connectivity Lemma"): Let B_1 and B_2 be two optimal bases solving, for fixed $L > 0$, the problem Max LCx subject to x in F. There exists a finite sequence of neighboring bases--whose tableau meet the efficiency criterion--linking

the two.

Proof: It is well known that the simplex may be used to find all alternate optima to a linear program. In so doing, it passes from basis to neighboring basis (by pivoting according to the exit criteria (a) and (b) above), each of which satisfies the linear programming optimality criterion--and hence the efficiency criterion. Suppose that the process begins at B_1; eventually B_2 is found. Tracing backward along the path of discovery yield the sequence sought.

QED

The following rather pedestrian identity is required in later proofs. Parametric objective functions will figure in these proofs, but for the moment it is sufficient to understand that V' and V'' are two objective gradients (that is, $V'x$ and $V''x$ are objective functions) and a third, $V(s)$, is given by $V(s) = V' + s(V''-V')$ where s is a scalar.

Lemma V-8 ("Reduced Costs Identity"): Let V' and V'' be two objective gradients and let $V(s) = V' + s(V''-V')$. Let $g(j,i,s)$ be the reduced cost for the variable x_j corresponding to objective $V(s)$ and basis B_i, and let $g(j,i)$ be the reduced cost for that variable corresponding to basis B_i and the objective $(V''-V')x$. Then for any given scalar s^*,

$$g(j,i,s) = g(j,i,s^*) + (s-s^*)g(j,i)$$

Proof: By the definition of reduced cost,

[illegible]

[illegible]

[illegible]

$$g(j,i,s) = V(s)_j - V(s)_{B_i} B_i^{-1} a_{.j}$$

where $V(s)_{B_i}$ is the vector of components of $V(s)$ corresponding to the variables x_j which are basic at B_i (and recall that $a_{.j}$ is the j^{th} column of A). Since $V(s) = V' + s(V''-V')$ this expression can be written:

$$g(j,i,s) = V'_j + s(V''_j - V'_j) - (V'_{B_i} + s(V''_{B_i} - V'_{B_i}))B_i^{-1} a_{.j}$$
$$V'_j - V'_{B_i} B_i^{-1} a_{.j}) + s \left\{ (V''_j - V'_j) - (V''_{B_i} - V'_{B_i}) B_i^{-1} a_{.j} \right\}$$

Substituting in the identity $s = s^* + (s-s^*)$ gives:

$$g(j,i,s) = (V'_j - V'_{B_i} B_i^{-1} a_{.j}) + s^*[(V''_j - V'_j) - (V''_{B_i} - V'_{B_i}) B_i^{-1} a_{.j}]$$
$$+ (s-s^*)[(V''_j - V'_j) - (V''_{B_i} - V'_{B_i}) B_i^{-1} a_{.j}]$$
$$= g(j,i,s^*) + (s-s^*)\ [(V''_j - V'_j) - (V''_{B_i} - V'_{B_i}) B_i^{-1} a_j]$$
$$= g(j,i,s^*) + (s-s^*) g(j,i)$$

QED

Lemma V-9: $g(j,i,s)$ and $g(j,i)$ are defined as in the previous lemma. Let T be the set of all alternate bases B_i solving for fixed s (in the interval $0 \leq s \leq 1$) the problem:

$$\text{Max } V(s)x$$
$$\text{subject to } x \text{ in } F$$

Some basis B_i in the set T has one of the following properties:

(1) $g(j,i) \leq 0$ for all j

(2) $s' = \underset{j/g(j,i) > 0}{\text{minimum}} \quad -g(j,i,s)/g(j,i) > 0$

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

Proof: First it will be shown that there exists some basis in T which solves, for some positive e, the linear program:

$$\text{Max } V(s+e)x$$
$$\text{subject to } x \text{ in } F$$

Select arbitrary small e and solve the problem above. Let T' be the set of bases solving this problem. If some member of T' is also a member of T, proceed to the second part of the proof. Otherwise, let B_i be some member of T' and notice that (from the reduced cost identity):

$$g(j,i,s+e) = g(j,i,s)+e\ g(j,i) \text{ for all } j \qquad (*)$$

Since B_i solves the problem above, $g(j,i,s+e) \leq 0$ for all j. But since B_i is not in T (hence does not solve Max V(s)x for x in F) $g(j,i,s) > 0$ for at least one j. Find a new value of e as follows:

$$e' = 1/2 \underset{j/g(j,i,s) > 0}{\text{minimum}} \ (-g(j,i,s)/g(j,i))$$

$$\text{new } e = \text{minimum } (e', 1/2 \text{ of old } e)$$

e' is positive, because from (*), and the comment immediately following, when g(j,i,s) is positive, g(j,i) must be negative. Then e is positive also, and is smaller than the previous e. Resolve the linear program for the new value of e, and repeat the above analysis. Continue to repeat until some basis in T is found to solve the problem for some positive e. Such a basis will ultimately appear, for the following reasons. First, the boundedness of F guarantees

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

that each linear program has a solution. Secondly, e' is computed to be the smallest value of e such that the tableau associated with the associated basis satisfies the optimality criterion. Since the new e is smaller than e', and every succeeding e is smaller yet, the basis giving rise to e' can never again appear in a later T_i. But the number of bases is finite; eventually all bases not in T must be exhausted, and then the desired result must obtain.

(second part) Let B_i be the basis found which solves both

$$\text{Max } V(s+e)x \text{ subject to } x \text{ in } F$$

and

$$\text{Max } V(s) \text{ subject to } x \text{ in } F$$

The two optimality criteria this basis satisfies are:

$$g(j,i,s+e) = g(j,i,s) + e\, g(j,i) \leqq 0 \text{ for all } j$$

$$g(j,i,s) \leqq 0 \text{ for all } j$$

Together these imply that one of the following conditions must hold:

(1) $g(j,i) \leqq 0$ for all j

(2) $s' = \underset{j/g(j,i,s) > 0}{\text{minimum}} \; -g(j,i,s)/g(j,i)$ exists and is > 0.

QED

<u>Theorem V-10 ("Connectivity Theorem")</u>: If B' and B" are two different bases whose tableaux satisfy the efficiency criterion, then there exists a sequence of neighboring bases, whose

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

tableaux satisfy the efficiency criterion, linking the two.

Proof: Let A' and A'' and x' and x'' be the matrices of tight constraints and the vertices corresponding to B' and B''. By corollary V-3-2, there exist vectors k', k'', $L'>0$, and $L''>0$ such that $k'A' = L'C$ and $k''A'' = L''C$, and where when constraint i is an inequality in the statement of F then $k_i' \geqq 0$, or $k_i'' \geqq 0$, as appropriate. For some such k', k'', L', L'' define:

$$V' = L'C$$

$$V'' = L''C$$

If $V' = V''$ then B' and B'' are alternate optimal bases to

$$\text{Max } L''Cx \quad \text{subject to } x \text{ in } F$$

By the connectivity lemma (lemma V-7) the desired sequence exists. If $V' \neq V''$ define $V(s) = V' + s(V''-V')$ for scalar s. Notice that $V(0) = V'$, $V(1) = V''$, and

$V(s) = (sL'' + (1-s)L')C$. For s in the interval $0 \leqq s \leqq 1$, the vector multiplying C is strictly positive, so by theorem V-2 all solutions to "Max $V(s)x$ subject to x in F" for fixed s in that interval are efficient.

It is desired to find the bases which are optimal to the following linear program as s is increased from zero to one:

$$\text{Max } V(s)x \text{ subject to } x \text{ in } F$$

This is a well known single parameter parametric problem; the simplex may be employed to find the succession of optima which arise as s is increased.

[illegible]
[illegible]
[illegible]
[illegible]
[illegible]
[illegible]
[illegible]
[illegible]

$$[illegible]$$

[illegible]

[illegible]
[illegible]

[illegible]
[illegible]
[illegible]

[illegible]
[illegible]
[illegible]
[illegible]

[illegible]
[illegible]
[illegible]

[illegible]
[illegible]
[illegible]

The bases so generated constitute the desired sequence. In the following, the method of construction of this sequence is explained, and it is demonstrated that only efficient bases are included, and that B" must ultimately be reached. Define $g(j,i,s)$ and $g(j,i)$ as in the statement of the reduced costs identity.

Set $i = 1$, $B' = B_1$, and $s_1 = 0$ and commence the following:

(Iteration beginning)

B_i is efficient; therefore, $g(j,i,s_i) \leqq 0$ for all j. Recall that the reduced costs identity specifies:

$$g(j,i,s_{i+1}) = g(j,i,s_i) - (s_i - s_{i+1})g(j,i)$$

Several cases may arise:

Case (a): One of the following two conditions obtains:

(1) $g(j,i) \leqq 0$ for all j

(2) $s_{i+1} = \underset{j/g(j,i) > 0}{\text{minimum}} \left(-g(j,i,s_i)/g(j,i)\right) + s_i$ exists and is greater than or equal to one. (The determination of s_{i+1} in this fashion will be called the "subproblem").

In either of these events, B_i and B" are alternate optima to

$$\text{Max } V(s_i)x \text{ subject to } x \text{ in } F$$

By the connectivity lemma (lemma V-7) the desired sequence exists.

Case (b): The solution s_{i+1} to the subproblem above is $> s_i$ but less than 1. Then apply the exist criterion of the simplex to any problem attaining the subproblem minimum, and pivot as indicated

(the boundedness of F ensures that the exit criterion will be satisfied by some basic variable). Label the new basis B_{i+1}. This basis is feasible and, since it solves

$$\text{Max } V(s_{i+1})x \text{ subject to } x \text{ in } F$$

for s_{i+1} in the 0-1 interval, it is efficient. Of course, it neighbors B_i. Set $i = i+1$ and commence another iteration.

Case (c): The solution to the subproblem exists but $= s_i$. In this event, there are alternate optimal bases to the problem

$$\text{Max } V(s_i)x \text{ subject to } x \text{ in } F$$

By lemma V-9 at least one of these (call it B_t) has one of the following properties:

(1) $g(j,t) \leqq 0$ for all j

(2) $s_{i+1} = \underset{j/g(j,t)>0}{\text{minimum}} \; (-g(j,t,s_i)/g(j,t))$ exists and is > 0.

By the connectivity lemma (lemma V-7) this basis can be reached by a finite sequence of neighboring efficient bases. Once reached, the situation will fall into one of the categories in case (a) or (b) and the remarks there apply.

Notice that the sequence so constructed will terminate only when B" is reached. Moreover, the number of bases generated in the course of each iteration is finite. Also, s_i is increased each iteration in such a manner that the basis used in its construction can not be

optimal for succeeding s_i. Since the number of bases is finite, B" must ultimately appear.

QED

The following definitions and abbreviations will be employed in the statement of the algorithm for finding all efficient vertices to problem P and in Chapter VI.

Problem set AP: This is the family of problems which deter- the efficiency of the bases neighboring a given basis. Let G be the reduced costs matrix for the basis whose neighbors it is desired to evaluate, and suppose that the dimensions of G are k by q, where k is the number of rows of C and q is the number of nonbasic variables, including slack variables, in the constraints defining F. Let M be a k-vector, T a q-vector (with components T_i), I the rank q identity matrix, and 1 a vector of ones. Problem set AP is the following linear program for each i such that x_i is nonbasic:

$$\begin{aligned} &\text{Min } T_i \\ &\text{s.t. } G^t M + IT = - G^t 1 \\ &M,T = 0 \end{aligned}$$

LEB: List of efficient bases. (It is sufficient to keep only the indices of the basic columns). Bases are

[illegible]

indexed "LEV_i".

An Algorithm for Finding All Vertices Efficient to Problem P.

Step #1: Find an initial efficient basis; enter it on the LEB as LEB_1. Set $i = 1$, $j = 0$.

Step #2: Set $j = j+1$. If $j > i$, stop. Otherwise generate the tableau associated with LEB_j.

Step #3: Do the following for each nonbasic variable x_k:

(a) Ascertain which variables could leave the basis if the variable x_k were to enter. The exit criteria are:

$$d_t = 0 \text{ and } y_{tk} \neq 0, \text{ or}$$

$$d_t/y_{tk} = \min_{s/y_{s,k} > 0} (d_s/y_{sk})$$

(b) LEB_j is the list of the basic variables in the current basis. For each basic variable meeting the exit criteria above, form a new list of basic variables by substituting x_k for this exiting variable in the LEB_j. If at least one such new basis is not already listed on the LEB, solve the corresponding problem AP. If this problem has a solution, add the new unlisted bases to the LEB and increment i by 1 for each addition.

When all nonbasic variables have been so considered, go to step #2.

[illegible]

<u>Theorem V-11</u>: If x'' is an efficient vertex to problem P, then some basis representing x'' is on the LEB at the conclusion of the preceding algorithm.

Proof: Let B' be the initial efficient basis found in step #1. By theorem V-6, some basis representing efficient vertex x'' has a tableau satisfying the efficiency criterion. Let one such basis be B''. If $B'' \neq B'$ then by theorem V-10 there exists a sequence of neighboring bases, whose tableaux satisfy the efficiency criterion, linking the two. The neighbor by neighbor search of steps #2 and #3 uncovers each of these bases, including B''.

QED

V-5 AN ALGORITHM FOR FINDING THE EFFICIENT MANIFOLD

Not all efficient solutions to problem P need be vertices; finding expressions for those which are not is the task of this section.

Assume that all efficient vertices are known. It was remarked in section V-2 that facets may be expressed as convex combinations of vertices of F; clearly, then, efficient facets may be expressed as convex combinations of efficient vertices of F. A set of efficient facets may be expressed in terms of groups of efficient vertices where it is understood that any convex combination of the vertices in each group is also efficient. This form of expression may be used to specify the efficient manifold provided that all efficient solutions are contained in efficient facets. The following lemma and theorem establish that proviso.

Lemma V-12: If solution x is efficient to problem P then at least one constraint is tight at x.

Proof: If problem P contains equality constraints, the conclusion is immediate. If not, suppose that x is efficient but no constraint is tight at x. From the statement of problem P there does not exist row vector $L > 0$ such that $LC = 0$; therefore, by Steimke's theorem of the alternative, there exists x' such that $Cx' \geqq 0$. Let K be the set of all j such that $a_j x' > 0$ (recall that the coefficients of the nonnegativity constraints are included in A).

Let $e = \begin{cases} \underset{j \text{ in } K}{\text{minimum}} \left((b_j - a_j x)/a_j x' \right) & \text{if } K \text{ is not vacuous} \\ 1 & \text{if } K \text{ is empty} \end{cases}$

Let $x'' = x + ex'$

Notice that $a_i x'' = a_i x + e a_i x'$

There are two cases:

(a) $a_i x' \leqq 0$; then $a_i x'' \leqq b_i$ so x'' satisfies constraint i.

(b) $a_i x' > 0$;

$$a_i x'' = a_i x + \left(\underset{j \text{ in } K}{\text{minimum}} \left((b_j - a_j x) / a_j x' \right) \right) a_j x'$$

$$\leqq a_i x + ((b_i - a_i x)/a_i x') a_i x'$$

$$\leqq b_i$$

therefore constraint i is satisfied at x''.

Then x'' is in F.

But $Cx'' = Cx + eCx' \geqq Cx$ because $e > 0$, $Cx' \geqq 0$, and at least one component $(Cx')_i > 0$.

Then x is not efficient; a contradiction.

QED

Theorem V-13: If point x is efficient to problem P then x is contained in an efficient facet.

Proof: Let J be the set of indices of constraints tight at x (by the previous lemma, J is not vacuous). Consider the facet generated by the constraint set J; this facet is efficient (by corollary V-3-2) if there exist vector $L > 0$ and multipliers k_i (nonnegative when constraint i is an inequality)

such that

$$LC = \sum_{i \text{ in } J} a_i k_i$$

But since x is efficient, such L and k_i exist.

QED

Notice that every efficient vertex is some efficient facet.

It is appropriate at this point to comment on some other conclusions about facets. It is a natural first assumption that facets containing efficient vertices must be efficient. This is not true. Nor is it true that a facet every extreme point of which is efficient is necessarily efficient (a counterexample is included in Appendix II). On the other hand, a facet containing an inefficient vertex can not be efficient. This useful result will be taken up in theorem V-15. A most valuable implication of this is that the defining index set of every nonvacuous efficient facet must be contained in the defining index set of some efficient vertex.

The reader will recall that "facet" was defined in association with a collection of tight constraints. This suggests the following useful concept:

"Generic member": The "generic member" of a facet (or, occasionally, a "generic point" in a facet) will be defined to be any solution x satisfying as tight

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

constraints those and only those constraints which define the facet.

The generic member need not actually exist for all facets (that is, there may be no point in F actually satisfying those and only those constraints as equalities); nevertheless, whether it exists or not, it will be said to be efficient if (by corollary V-3-2) there exist $L \geqq E > 0$ and multipliers k_i ($\geqq 0$ when constraint i is an inequality) such that for the index set J of defining constraints

$$LC = \sum_{i \text{ in } J} a_i k_i$$

"Generically efficient": A facet will be said to be generically efficient if its generic member is efficient.

"Not generically efficient": A facet will be said to be not generically efficient if its generic member is not efficient.

In addition to generic members, facets may also contain other members (among them, perhaps, the vertices of the facets, which are also vertices of F) which satisfy additional constraints as tight constraints. Consider some such point satisfying the additional tight constraints indexed by the set J'. Clearly if the generic member is efficient, then this solution is also; to show this, it suffices to add $k_i = 0$ for i in J' to the L and k_i, i in J, satisfying the corollary V-3-2 condition for the generic point, to

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

get

$$LC = \sum_{i \text{ in } J} a_i k_i + \sum_{i \text{ in } J'} a_i k_i$$

which satisfies the efficiency conditions of that corollary for this other member as well. It follows that the efficiency of the generic member is sufficient to show the efficiency of the facet.

Testing the efficiency of the generic member can be cast into a linear programming format by the by now familiar device of substituting $M = L-1$ into the condition cited; the condition becomes

Find M_i, k_i, i in J, such that

$$\sum_{i \text{ in } J} MC - \sum_{i \text{ in } J} a_i k_i = -1C$$

$$M_i \geqq 0$$

$k_i \geqq 0$ when constraint i is an inequality constraint.

If no feasible solution to this point exists, then no generic member of the facet is efficient, although nothing can be said about other members of this facet. However--to preview theorem V-15--clearly there are no efficient solutions to P whose tight constraints are a subset of J.

In a problem with n + m variables and slack variables and m constraints, one can form $2^{n+m} -1$ different nonvacuous index sets. One way to find the efficient manifold is to find and test the

[illegible]

$$[illegible]$$

[illegible]

[illegible]

$$[illegible]$$

$$[illegible]$$

[illegible]

[illegible]

[illegible]

efficiency of each of the facets these index sets define. Having done so, the vacuous ones may be somehow weeded out, and the remainder represent the efficient manifold. The technique is simple, foolproof, and impractical for all but the smallest problems. Fortunately there are numerous shortcuts which can be employed. One, that only subsets of vertex defining index sets need be considered, was mentioned earlier. The following theorems motivate two more.

<u>Theorem V-14 (E - inference theorem)</u>: If J and J' are the constraint index sets defining two facets, if f(J) is efficient, and if J' includes J, then f(J') is efficient also.

Proof: The efficiency of the first facet implies there exist $L \doteq E > 0$ and k_i (nonnegative for inequality constraints) such that

$$LC = \sum_{i \text{ in } J} a_i k_i$$

Then

$$LC = \sum_{i \text{ in } J} a_i k_i + \sum_{\substack{i \text{ in } J' \\ i \text{ not in } J}} a_i k_i$$

for $k_i = 0$, i in J' but not in J, so the generic member of f(J') and hence f(J') itself are efficient.

QED

Theorem V-15 (NGE-inference theorem): If J and J' are the defining index sets of f(J) and f(J'), if J includes J', and if f(J) is not generically efficient, then f(J') is not generically efficient.

Proof: If there does not exist $L > 0$ and k_i ($\geqq 0$ when constraint i is an inequality) such that

$$LC = \sum_{i \text{ in } J \text{ and } J'} a_i k_i + \sum_{\substack{i \text{ in } J \\ i \text{ not in } J'}} a_i k_i$$

then there can not exist $L > 0$ and k_i (restricted as before) such that

$$LC = \sum_{i \text{ in } J} a_i k_i$$

QED

Fortunately, the E-inference theorem and the NGE-inference theorem can be used to considerably shorten the facet search procedure outlined earlier. To begin with, the reader should be aware that it is not generally necessary to list all efficient facets in order to specify the efficient manifold. This is because often some efficient facets are contained in others (this is the case when the defining index set of the second contains that of the first). Whenever such cases occur, listing the large facet suffices to identify not only all efficient solutions contained in it, but also those

contained in the smaller facet as well. It follows that in the course of searching for efficient facets, when one is found, its including facets (that is, those defining index sets containing its defining index set) can safely be ignored. Conversely, when a not generically efficient facet is found, those facets including it (that is, whose defining index sets are included in the index set of the first facet) can be ignored. This is the principle alluded to earlier, when it was noted that only defining index sets contained in the defining index set of some efficient vertex need be considered.

Unfortunately these twin principles work in opposite directions. When one is looking at facets of one constraint, the E-inference theorem is potentially quite valuable, because the opportunity for these facets to include others is high. On the other hand, the NGE inference theorem is worthless; no nonvacuous proper subsets of the list of tight constraints exist, so no efficient facets can contain these. If one's facet search progresses from large facets (with few defining indices) toward smaller constraints (with many defining indices) then the NGE-inference theorem remains useless, for all it can tell is what has already been painstakingly discovered--that larger including facets are not generically efficient. Conversely, a facet search which progresses from large index set facets toward small index set facets finds no value in the E-inference theorem. Starting with middlesized facets reduces the potential of both theorems. The upshot is that there

seems to be no a priori grounds for preferring any particular facet search above all others. The choice for a particular application must revolve about practical considerations including the nature of the problem and the capabilities of the computer at hand. The efficient facet search procedure about to be presented is motivated by the anticipated virtue of a "self fathoming" feature (an opportunity to avoid useless computations by structuring the sequence of operations to take full advantage of the E-inference theorem) which it exhibits.

The following definitions and abbreviations will be employed in the statement of the algorithm (and in the following chapter):

"Problem Q" This is the problem which determines whether the generic member of a facet (hence the facet itself) is efficient. Let M be a k-vector, 1 a vector of ones, and k_i scalar multipliers (nonnegative whenever constraint i is an inequality constraint). "The problem Q associated with set J" (or with facet f(J)) is to find M and the k_i such that

$$MC - \sum_{i \text{ in } J} a_i k_i = -1C$$

$$M = 0$$

$k_i \geqq 0$ when constraint i is an inequality constraint.

(This is a problem for phase one of the two phase simplex).

[illegible]

LEV: List of efficient vertices. The information stored for each vertex is the indices of the constraints which are tight at that vertex. LEV_i indicates the i^{th} vertex stored on the LEV.

The LEV is easily constructed from the complete LEB by ascertaining for each member of the LEB which constraints are tight at the vertex represented. The resulting vertex list is then checked for duplication (which will arise whenever several bases from a single degenerate vertex appear on the LEB).

LEF: List of efficient facets. The information stored for each facet is its defining index set. LEF_i indicates the i^{th} facet on the LEF.

LCF: List of candidate facets. Facets are specified by their defining index sets; LCF_i is the i^{th} facet on this list. The function of this list will become clear once the algorithm is understood.

An Efficient Facet Finding Algorithm

Step #1: Store the defining index sets of the efficient vertices in LEV_i, $i = 1,2,\ldots,U$. Set S = the set of indices of equality constraints. Attempt to solve the problem Q associated with S. If a solution exists, place S on the LEF and stop. Otherwise set i = 0.

Step #2: Set $i = i+1$. If $i > U$ stop. Otherwise set $J' = S$; set J'' = the defining index set of LEV_i $= (S, i_1, i_2, \ldots, i_r)$. Set LCF = the empty set.

Step #3: Set i_u = the index with the smallest subscript in $J'' - S$ which is greater than all index subscripts in $J' - S$. If there is no such index, go to step #6. Otherwise set $k = u$.

Step #4: Form the index set $J' + i_k$. Determine whether this index set contains any index set on the LEF; if so, go to step #5. If not, attempt to solve the associated problem Q. If a solution exists, place this index set on the LEF; otherwise place it at the bottom of the LCF.

Step #5: Set $k = k + 1$. If $k > r$ go to step #6; otherwise go to step #4.

Step #6: If the LCF is empty, go to step #2. Otherwise remove the top index set from the LCF; call this set J'. Go to step #3.

This algorithm is clearly finite because the number of vertices is finite, the number of subsets of the defining index set of any vertex is finite, and no subset is generated more than once in the course of investigating the defining index set of a given vertex.

Example V-1:

Suppose that LEV_1 of a given problem has the defining index set (S,i_1,i_2,i_3,i_4) where i_j indexes the j^{th} constraint, and suppose that of the LEF for this problem, two facets ($f(S,i_1,i_3)$ and $f(S,i_4)$) have index sets contained in the defining index set of LEV_1. Then, substituting the values of the subscripts for the constraint indices --e.g. "3" signifies i_3, and representing the problem Q associated with index set J by Q(J), the algorithm processes the defining index sets arising from this vertex as follows:

Q(S) is found not to have a solution.

Q(S,1), Q(S,2), Q(S,3) are found not to have solutions; the sets (S,1),(S,2),(S,3) are placed on the LCF.

Q(S,4) has a solution; (S,4) is placed on the LEF.

k now exceeds r (which is 4) so the top member of the LCF, (S,1) is removed (and k is reset).

Q(S,1,2) is found to have no solution; (S,1,2) goes to the bottom of the LCF.

Q(S,1,3) has a solution; (S,1,3) is placed on the LEF.

Set (S,1,4) is generated but found to contain (S,4) from the LEF; it is discarded.

(S,2) is removed from the LCF.

Q(S,2,3) has no solution; (S,2,3) is placed at the bottom of the LCF.

Set (S,2,4) is generated but found to contain (S,4).

(S,3) is removed from the LCF; (S,3,4) is generated but found to contain (S,4).

[illegible]

(S,1,2) is removed from the LCF; both (S,1,2,3) and (S,1,2, 4) are found to contain members of the LEF.

(S,2,3) is removed from the LCF; (S,2,3,4) contains a set on the LEF.

The LCF is now empty; the algorithm passes to LEV_2. The relevant portion of the LEF has indeed been found; moreover, while there are 15 possible subsets of LEV_1, only 8 attempted solutions of problems Q were required.

Theorem V-16: Solution x is efficient to problem P if and only if x is contained in some facet on the LEF constructed by the preceding algorithm.

Proof: (If clause): A facet is placed on the LEF (in step #4) only if its corresponding problem Q has a solution; this is sufficient for the efficiency of all solutions contained in the facet. (Only if clause): Let x be efficient to problem P. By theorem V-13 x is contained in some efficient facet. Let J" be the defining index set of one such facet and let J be the smallest subset of J" (there may be several--any one will do) containing S (the indices of the equality constraints) and defining an efficient facet. If J = S then J is placed on the LEF in step #1. Otherwise, note that since f(J) contains x, it contains at least one vertex on the LEV. Consider the first such vertex listed there; its defining index set contains J. In due course this vertex is taken up

for consideration in step #2. In step #4 thereafter a set containing S and the inequality constraint index in J with the lowest subscript is formed. Since this is the first vertex containing J, J (and of course any subset) can not already be on the LEF. If this index set is J, then the corresponding problem Q is solved and J is placed on the LEF. Otherwise no solution to problem Q exists (guaranteed by the method of construction of J) so the index set is placed on the LCF. In due course it is recovered and the inequality index of next lowest subscript is added. If this is J, it is placed on the LEF; otherwise, it is placed on the LCF. The process continues at least until J is constructed in its entirety and placed on the LEF.

QED

The facet search procedure just outlined is reasonably easy to understand and straightforward to program, yet it does exhibit some computational niceties. For example, no problem Q associated with an efficient facet is solved twice; after its first appearance, the facet is listed on the LEF, which is checked before each problem Q is attempted. Moreover, no facets contained in other efficient facets are investigated; this is the "self fathoming" feature. On the other hand, storing the LCF may sometimes be a considerable burden, and as written, nothing prevents repeated attempts to find solutions to problems Q corresponding to facets which are not

generically efficient.

It is possible, by introducing a slight complication, to remedy the last mentioned fault. In order to accomplish this, an additional operation must be inserted in step #4 between the LEF check and the attempt to solve problem Q. It is desirable at this point to check to see if the index set under consideration is contained in the defining index set of some previously investigated member of the LEV. If this is the case, then place the set on the LCF and go to step #5. In effect, the set has been inferred to represent a not generically efficient facet, for if it were efficient, it or a subset would have appeared on the LEF during consideration of a previous LEV. With this adaptation, the algorithm's handling of problems with many efficient vertices is expedited considerably.

The goal of this chapter has been to develop a way to express the efficient manifold. It is now possible to do so. The efficient manifold is composed of one or more efficient facets; each of these may be expressed as the set of all convex combinations of its vertices. Vertices are easily associated with the facets containing them; a vertex is in a facet if the defining index set of the former contains that of the latter. In addition to being specified in terms of defining index sets, each vertex has been identified as a solution "x". This information was required for application of the exit criteria in the efficient vertex search, and for the conversion of the LEB into the LEV. Defining the vertices of

[illegible]

[illegible]

each facet in terms of the solutions x which they represent accomplishes a precise specification of the efficient manifold, the set of all efficient solutions to problem P.

FOOTNOTES

[1]Simmonard, M. Linear Programming, Prentice-Hall, Englewood Cliffs, New Jersey, 1966.

[2]Mangasarian, O. L. Nonlinear Programming, McGraw-Hill, New York, 1969.

[3]Philip, J. "Algorithms for the Vector Maximization Problem," Mathematical Programming, II (1972), pp. 207-229.

CHAPTER VI

ALTERNATIVE FACET FINDING TECHNIQUES

VI-1 GENERAL

The previous chapter introduced a method for finding an expression for the efficient manifold. That method was comprised of two phases; an algorithm for finding the set of efficient vertices, and another for finding efficient facets. The "phase two" algorithm already presented is but one of many ways to go about finding these efficient facets. Three significant characteristics of the application of a given facet search procedure to a given problem are the number of problems Q attempted, the greatest lengths attained by the various lists kept in conjunction with the operation of the algorithm, and the complexity of the algorithm itself. These constitute measures of the running time of the algorithm, its storage requirements, and its computer coding complexity. Though it is desirable to minimize all three, they are to some degree inherently conflicting. The importance to be attached to any one of them will vary with the circumstances of the user and the character of the problem at hand. There is no ideal facet search procedure apparent.

This chapter discusses three alternative ways to carry out the second phase of the efficient manifold presentation method. The three approaches are:

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

(a) A facet search algorithm similar to phase two of the method of Chapter V, except that the search progresses from facets of many constraints to facets of few constraints. This algorithm (the "bottoms-up" algorithm) is presented in section VI-2.

(b) A facet search procedure combining the approaches of (a) above and of the second phase of the Chapter V approach. This algorithm (the "double-ended" algorithm) is presented in section VI-3.

(c) An approach to facet finding which requires less information to be stored in the course of the algorithm but which can be expected to require more subproblems (linear programs) to be solved than any other facet search procedure discussed. This approach (the "low storage algorithms") is presented in section VI-4.

VI-2 A "BOTTOMS-UP" EFFICIENT FACET FINDING ALGORITHM

In the "bottoms-up" algorithm, the efficient facet search begins by considering facets with large numbers of tight constraints, and works toward those with few such. As in all of the facet search procedures, the search is founded upon the tight constraint sets of the efficient vertices--which, the reader will recall, are facets themselves. The following names and abbreviations will be used in the statement of this algorithm:

LEV: The list of efficient vertices, each specified by the list of indices of the constraints which are satisfied as equalities at the vertex in question.

LCF: A list of candidate facets, specified by lists of their tight constraint indices.

LNGF: A list of not generically efficient facets, similarly specified.

LEF: The list of efficient facets, similarly specified.

Q(J): The problem Q associated with the constraint index set J.

Defining index set: A set of constraint indices. The defining index set corresponding to a vertex or facet is the set of indices of constraints satisfied as equalities.

The members of each list will be designated by subscripting the symbol for the list--e.g., LEV_k is the defining index set representing the k^{th} vertex on the list LEV.

The function of the variable "flag" in the bottoms-up and double-ended algorithms is to keep track of whether or not a given defining index set under consideration has given rise to another index set, a subset of the first, which is stored on the list of candidate facets. If not, flag = 0 is the signal for the index set to be placed on the LEF.

The bottoms-up algorithm:

Step #1: Store the defining index sets of the efficient vertices in LEV_i, $i = 1,2,\ldots,U$. Set S = the set of indices of the equality constraints. Attempt to solve the problem Q associated with S. If a solution exists, place S on the LEF and stop. Otherwise set $i = 0$.

Step #2: Set $i = i+1$. If $i > U$, stop. Otherwise set J'' = the defining index set of $LEV_i = (S, i_1, i_2, \ldots, i_r)$. Set $J = J''$. Set LCF and LNGF = the empty set.

Step #3: Set variable "flag" = 0. Find i_u, the inequality constraint index with the lowest subscript in J which is higher than all indices in $J'' - J$. If no such i_u exists, go to step #6. Otherwise set $k = u$.

Step #4: Form index set $J' = J - i_k$. If J' is contained in some LEV_j, $j < i$, or contained in some $LNGF_j$, go to step #5. If J' contains LEF_j set flag = 1, place J' at the bottom of the LCF, and go to step #5. Otherwise attempt to solve problem $Q(J')$. If a solution exists,

set flag = 1 and place J' at the bottom of the LCF. Otherwise add J' to the LNGF (remove from the LNGF any sets contained in J').

Step #5: Set $k = k + 1$. If $k \leq r$, go to step #4.

Step #6: If flag = 0 and if J contains no set in the LEF, add J to the LEF and remove sets containing J. If the LCF is empty, go to step #2. Otherwise remove the topmost member of the LCF; call this set J; go to step #3.

Example VI-1:

In this example the bottoms-up algorithm is applied to the problem of Example V-1. Recall that LEV_1 has the defining index set (S,i_1,i_2,i_3,i_4) and that the relevant portion of the LEF is $f(S,i_1,i_3)$ and $f(S,i_4)$. Hereafter, subscripted indices will be represented by their subscripts alone; Q(J) indicates the problem Q associated with the set J; "Yes" will mean that this problem has a solution; "No" indicates otherwise. "in" indicates either that the set is a subset or superset of the following list; whether subset or superset will be clear from the context. "to" means the set is to be placed on the list indicated. When this is the LCF, placement at the bottom is understood. For example "(S,1,4) is in LEF" says that the set (S,i_1,i_4) contains some set on the LEF; "Q(S,4) Yes; to LCF" means that the problem Q associated with the set (S,i_4) has been attempted and a solution is found, and that the set (S,i_4) is to be placed on the LCF, at the bottom. Only the sets and problems Q

arising in step #4 of the algorithm will be listed completely; other information will be supplemented where clarity demands. The algorithm proceeds as follows:

Q(S) No

flag = 0

Q(S,2,3,4) Yes: to LCF; flag = 1

Q(S,1,3,4) Yes: to LCF

Q(S,1,2,4) Yes: to LCF

Q(S,1,2,3) Yes: to LCF

The LCF yields set J = (S,2,3,4); flag = 0

Q(S,3,4) Yes: to LCF: flag = 1

Q(S,2,4) Yes: to LCF

Q(S,2,3) No: to LNGF

The LCF yields set J = (S,1,3,4); flag = 0

Q(S,1,4) Yes: to LCF; flag = 1

Q(S,1,3) Yes: to LCF

The LCF yields set J = (S,1,2,4): flag = 0

Q(S,1,2) No: to LNGF

flag is still = 0 when k exceeds r: (S,1,2,4) to LEF

LCF yields set J = (S,1,2,3): flag remains zero: to LEF

LCF yields set J = (S,3,4): flag = 0

Q(S,4) Yes: to LCF: flag = 1

(S,3) is in LNGF

LCF yields J = (S,2,4): flag reset = 0.

(S,2) in LNGF

flag remains = 0 when k exceeds r: (S,2,4) to the LEF:

remove (S,1,2,4) from the LEF

LCF yields J = (S,1,4): reset flag = 0

(S,1) is in LNGF: flag = 0 and k exceeds r: (S,1,4) to LEF

LCF yields J = (S,1,3): flag = 0

(S,1) is in LNGF: flag remains = 0: (S,1,3) to LEF (and

remove (S,1,2,3))

LCF yields J = (S,4): reset flag = 0

(S) is in LNGF: flag is still = 0: (S,4) to LEF (and

remove (S,2,4))

The LCF is now empty: the algorithm proceeds to consider LEV_2 etc. Twelve different problems Q were attempted here (none more than once).

It will now be shown that the bottoms-up algorithm finds all efficient facets.

Consider I, a set containing the equality constraint indices and an arbitrary subset of the inequality constraint indices of the defining index set of some LEV_m. Let J be the complement of I with respect to LEV_m, where $J = (j_1, j_2, \ldots, j_e)$ and $j_n > j_{n+1}$ for all n. Define the following sequence of index sets:

$J_1 = I + j_1 + j_2 + \ldots + j_e$ (= the index set of LEV_m)

$J_2 = I + j_1 + j_2 + \ldots + j_{e-1}$

$\ldots$

$J_e = I + j_1$

$J_{e+1} = I$

This sequence will be called the "discovery sequence" for I from LEV_m, because it is the sequence of sets leading to set I in the course of investigating the subsets of the defining index set of LEV_m. That is, suppose that I is generated in step #4. This set can only have been generated from the set J_e, by the following reasoning. I is generated from some set by the elimination of one member; then this predecessor set must contain I and one other element--that is, one element of J. Suppose that element were something other than j_1. Then j_1 is not in this prior set, and the extra element which is, is less than j_1. But since only indices greater than or equal to i_u may be eliminated, and since i_u is greater than all indices of the LEV which are missing from the set in question (step #3), and since the extra element is less than the missing j_1, that element can not be eliminated in step #4. By the same reasoning, J_e must have been generated from J_{e-1}, and so on, until J_1, the vertex's index set, is reached.

Remark: The discovery sequence for set I is unique because there is only one increasing order of the set J. It follows that the bottoms-up algorithm generates each set I no more than once --thus it attempts to solve problem Q(I) at most once.

Lemma VI-1: If set I is placed on the LCF, then from that time forward I or a subset will be listed on the LCF or LEF or be under consideration as the set J in steps #3 to #6.

Proof: When a set J on the LCF is removed, in step #6, then in the steps leading up to and including the next return to step

[illegible]

#6, either some subset of J is placed on the LCF (flag = 1) or else (flag remaining = 0) J is placed on the LEF. Sets on the LEF are removed only when proper subsets of them are placed thereon.

QED

Lemma VI-2: Let LEV_m be the first efficient vertex whose defining index set contains set I, which contains S, the set of equality constraint indices. Problem Q(I) has a solution if and only if I or some subset of I is contained in the LEF when the bottoms-up algorithm concludes the consideration of the subsets of the defining index set of LEV_m.

Proof: (If clause): Sets are placed on the LEF only after their problems Q are shown to have solutions.

(Only if clause): Assume that problem Q(I) has a solution. It will be shown that I or a subset of I must be on the LEF by the conclusion of consideration of LEV_m.

Consider the sets $J_1, J_2, \ldots,$ in the discovery sequence for I from LEV_m. Since I is a subset of each, no prior LEV contains any of them. Also, by theorem V-14, solutions exist to each of their problems Q. In step #1, Q(S) is tested and if a solution exists, S is placed on the LEF: the desired result is obtained. Otherwise the algorithm performs the following operations (among others) when LEV_m is reached.

J_2 is generated and since it represents an efficient facet, it is placed on the LCF before step #6 is reached. Eventually J_2 is removed from the LCF, J_3 is generated, and placed on the LCF, again before step #6 is reached. The algorithm proceeds to generate this sequence, some member of which is on the LCF whenever step #6 is reached, at least until the original set I is reached and considered. Of course, I, too, is placed on the LCF. By lemma VI-1, either I or a subset must be on the LEF at the conclusion of consideration of LEV_m, for the LCF is then empty.

QED

Theorem VI-3: Solution x is efficient to problem P if and only if x is contained in some facet on the final LEF constructed by the bottoms-up algorithm.

Proof: (If clause): Clearly x is efficient if it is contained in a facet on the LEF, for facets are listed thereon only if their problems Q have been shown to have solutions.

(Only if clause): Assume that x is efficient. By theorem V-13, x is contained in some efficient facet f(J'). This facet is nonvacuous (it contains x) so J' is contained in the defining index set of at least one efficient vertex. Let LEV_m be the first such on the LEV. By the previous lemma, J' or some subset is on the LEF at the conclusion of the algorithm. Let this set be J. Facet f(J) contains

$f(J')$; x is in $f(J)$.

QED

The bottoms-up algorithm can be simplified considerably, at some cost to its efficiency, by eliminating all reference to the LNGF, in which case there may occur more than one attempt to solve problems Q for which solutions do not exist, and by simplifying the rules of formation of the sets J' in steps #3 and #4 to be simply "form all sets which may be formed by the omission of a single inequality constraint index in J", in which case some redundancy will occur on the LCF (in effect, the discovery sequence for a given set will no longer be unique).

An embellishment which might pay for itself by reducing storage requirements would be to periodically purge from the LNGF any subsets of other sets on the LNGF. A convenient time to do this is when new sets are added to that list.

VI-3 A DOUBLE-ENDED FACET FINDING ALGORITHM

The algorithm described here is a simple combination of the algorithms of the previous section and of section V-5. The facets containing each efficient vertex are investigated starting with both the largest (few defining indices) and the smallest (many defining indices) and concluding with facets of an intermediate dimension. The two halves of the algorithm can profit from one another by the mechanisms of theorems V-14 and V-15. This appealing feature is gained at the expense of considerable added complexity, including two lists of candidate facets. The two are:

LCF1: The list of candidate facets used in the 1st portion of the algorithm, the part patterned after the algorithm of section V-5. Members of this list are not generically efficient.

LCF2: The list of candidate facets used in the 2nd portion --the bottoms-up portion of the algorithm. These facets are generically efficient.

With the exception of the LCF, the other lists defined heretofore are required in this algorithm also. List members are indexed and specified as in the previous algorithms. A new function and two additional variables are also required.

$c(J)$ is defined to be the number of indices--the cardinality --of the set J.

c' and c'' are the cardinalities of the most recently investigated members of LCF1 and LCF2.

The double-ended algorithm progresses through the possibilities at each efficient vertex by checking all candidate facets of cardinality one, then all candidates of the largest possible cardinality, then all facets of cardinality two, then all of cardinality one less than the largest, and so on. When $c' + 1 = c''$ then all possibilities of any cardinality have been investigated; the algorithm moves on to the next efficient vertex.

The Double-Ended Algorithm:

Step #1: Attempt to solve the problem Q associated with the equality index set S; if a solution exists, store S on the LEF and stop. Otherwise store the defining index sets of the efficient vertices in LEV_i, $i = 1,2,\ldots,U$. Then set $i = 0$.

Step #2: Set $i = i + 1$. If $i > U$, stop. Otherwise set LCF1=LCF2=LNGF= the empty set. Set J'' = the defining index set of $LEV_i = (S, i_1, \ldots, i_r)$ and let $J = S$, $J''' = J''$, and $c'' = c(J'')$

Step #3: Set $c' = c(J)$. Find i_u = the index in J'' with the smallest subscript greater than all index subscripts in J-S. If i_u does not exist go to step #6. Otherwise set $k = u$.

Step #4: Form the set $J' = J + i_k$. If J' contains some index set on the LEF, go to step #5. If J' is contained in some prior LEV_j or in some set on the LNGF, add J' to

the bottom of LCF1 and go to step #5. Otherwise attempt to solve the associated problem Q. If a solution exists, add* J' to the LEF; otherwise add it to the LCF1.

Step #5: Set $k = k+1$. If $k \leq r$, go to step #4.

Step #6: If LCF1 is empty, go to step #11. Otherwise remove the top member of that list; call it J. If $c(J) = c'$ go to step #3. Otherwise set $c' = c(J)$. If $c(J)+1 = c''$ go to step #11.

Step #7: Set $c'' = c(J''')$, flag = 0, and find i_u = the index in J''' with the smallest subscript greater than all index subscripts in $J'' - J'''$. If no such i_u exists go to step #10. Otherwise set $k = u$.

Step #8: Form the set $K = J'''-i_k$. If set K is contained in some prior member of the LEV or in some set on the LNGF then go to step #9. If K contains some index set on the LEF, set flag = 1 and place K at the bottom of the LCF2. Otherwise attempt to solve Q(K). If a solution exists set flag = 1 and place K at the bottom of LCF2. Otherwise add K to the LNGF.

Step #9: Set $k = k+1$. If $k \leq r$ go to step #8.

Step #10: If flag = 0 place* set K on the LEF. If LCF2 is

---

*If a set to be placed on the LEF contains a set already thereon, the new set should be ignored. Sets containing a newly added set should be removed from the LEF.

empty go to step #11. Otherwise remove the top member of the LCF2; call it J"'. If c(J"') = c" go to step #7. Otherwise set c" = c(J"'). If c" > c'+1 go to step #3.

Step #11: Place the remaining members of the LCF2 on the LEF. Go to step #2.

Example VI-2: Let LEV_1 = (S,1,2,3,4,5) and let f(S,2) be the only efficient facet containing LEV_1. The notation of the previous examples will be used. The double ended algorithm proceeds as follows.

Q(S) No

Q(S,1) No: to LCF1

Q(S,2) Yes: to LEF

Q(S,3) No: to LCF1

Q(S,4) No: to LCF1

Q(S,5) No: to LCF1

Now the bottoms-up portion of the algorithm commences:

(S,2,3,4,5) is in LEF: to LCF2

Q(S,1,3,4,5) No: to LNGF

(S,1,2,4,5) is in LEF: to LCF2

(S,1,2,3,5) is in LEF: to LCF2

(S,1,2,3,4) is in LEF: to LCF2

Now the Chapter V portion, again:

J = (S,1) from LCF1. (S,1,2) is in LEF

(S,1,3) is in LNGF: to LCF1

(S,1,4) is in LNGF: to LCF1

(S,1,5) is in LNGF: to LCF1

J = (S,3) from LCF1: (S,3,4) and (S,3,5) are in LNGF: both to LCF1

J = (S,4) from LCF1: (S,4,5) is in LNGF: to LCF1

J = (S,5); no additional set is generated from this set.

The algorithm returns to the bottoms-up portion of the algorithm:

J"' = (S,2,3,4,5), from LCF2. (S,3,4,5) is in LNGF

(S,2,4,5), (S,2,3,5), and (S,2,3,4) are all in LEF: to LCF2

J"' = (S,1,2,4,5), from LCF2. (S,1,2,5) and (S,1,2,4) are in LEF: to LCF2

J"' = (S,1,2,3,5) from LCF2. (S,1,2,3) is in LEF: to LCF2

J"' = (S,1,2,3,4) from LCF2, but no subsets of this are generated.

J"' = (S,2,4,5) and it is found that c(J"') = c'; the algorithm proceeds to step 11, finds that no additions to the LEF are required, and passes on to the next member of the LEV.

Only seven problems Q were attempted in the course of this solution, thanks to the early placement of an entry on the LEF and another on the LNGF. Either the bottoms-up algorithm or the algorithm of Chapter V would attempt seventeen problems Q in the course of investigating this efficient facet.

This algorithm is clearly finite, for the following reasons. The number of vertices is finite, and the progression of the algorithm does not repeat vertices. The number of subsets of the defining index set of any particular vertex is also finite. There are two enumeration procedures at work in the algorithm and neither repeats itself or the other; no subset of a particular vertex defining index set is generated more than once in the process of investigating that vertex. Then the consideration of each vertex is finite; the number of vertex investigations is finite; the algorithm is finite.

The discovery sequence discussed in conjunction with the bottoms-up algorithm is pertinent to this algorithm also, in conjunction with its bottoms-up portion, steps #6 - #9. Analagous to it is another discovery sequence operative during that portion of the algorithm patterned after Chapter V. This "downward discovery sequence" is constituted as follows. Let I be the set for which the sequence is to be constructed, and let it be composed of indices $i_1, i_2, \ldots, i_p$ where $i_j < i_{j+1}$. The enumeration process in the Chapter V portion of the algorithm generates the following sequence:

$$K_1 = i_1$$

$$K_2 = i_1, i_2$$

$$K_3 = i_1, i_2, i_3$$

$$\ldots$$

$$K_p = i_1, i_2, \ldots, i_p = I$$

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

Like the other sequence, once this downward discovery sequence is embarked upon, there will always be one member of the sequence under consideration or on the list of candidate facets, in this case the LCF1, until some discovery implying the efficiency (in this case) or not generic efficiency (in the case of the downward discovery sequence) of the remainder of the sequence is made. The event which can terminate this sequence is the discovery of the efficiency of some facet in the sequence. It is immaterial whether the efficiency of the member is itself determined directly or whether the efficiency of the facet generated by some subset of a member of the sequence is established in conjunction with another portion of the algorithm.

Theorem VI-4: Solution x is efficient to problem P if and only if x is contained in some facet on the LEF at the conclusion of the double-ended algorithm.

Proof: (If clause): If x is contained in a facet listed on the LEF then x is efficient, for facets are placed thereon only after their problems Q are shown to have solutions.

(Only if clause): Let x be efficient to problem P. By theorem V-13, x is contained in some efficient facet. This facet is nonvacuous, since it contains x. Its defining index set J will be contained in the defining index set of some efficient vertex. Consider the situation at the end of consideration of the first such vertex on the LEV.

Case (a): $c(J) \leqq c'$. Consider the downward discovery sequence for J. Let J' be the first set thereon which defines a generically efficient facet. J contains J' so x is contained in f(J'). Since no member of the sequence prior to J' defines a generically efficient facet, each has generated its successor, which was placed on the LCF1. Ultimately J' was generated, found to describe an efficient facet, and placed on the LEF.

Case (b): $c(J) \geqq c''$. Consider the discovery sequence for the set J. Every member of this sequence describes an efficient facet. Then every member has caused its successor to be placed on the LCF2. Then J can not fail to have been generated. Once placed on the LCF2 either it was placed on the LEF or a subset was placed thereon, either because no further efficient subset was generated, or because the set or subset was on the LCF2 when $c'+1 = c''$ obtained.

Case (c): $c' < c(J) < c''$ and the algorithm passed into step #11 when LCF1 was found to be empty in step #6. The remarks pertaining to case (a) apply; LCF1 can not become empty until some subset of J defining an efficient facet is found.

Case (d): $c' < c(J) < c''$ and the algorithm passed into step #11 when LCF2 was found to be empty in step #10. The remarks of case (b) apply; LCF2 can not become empty until J or a subset is found to be generically efficient. Then c(J)

must be $\leqq$ c" - a contradiction; case (d) can not arise.

When a set has been placed on the LEF, it is subsequently removed only if replaced by a subset of itself. Then at the conclusion of the algorithm some set containing solution x is on the LEF.

QED

Notice that during the course of the double-ended algorithm, problem Q(J) for arbitrary J is attempted at most once. This may be seen from the following reasoning. Other than Q(S), problems Q are attempted only in steps #4 and #8. In each case the LEV is checked prior to the attempt to solve the problem, so a given problem can arise during the investigation of at most one vertex. Now consider the situation at the end of consideration of a vertex in which Q(J) is investigated. Either $c(J) \leqq c'$ or $c(J) \geqq c''$ but not both; then set J can not have arisen in both the Chapter V and bottoms-up portions of the algorithm. The enumeration procedure within each of these halves of the algorithm does not generate a given set more than once.

This algorithm (and in fact any algorithm using the LNGF) could begin with an initial list of not generically efficient index sets in the LNGF if the not generically efficient bases found in the course of the efficient vertex search are saved for this purpose.

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

VI-4 A LOW STORAGE EFFICIENT FACET FINDING PROCEDURE

In the efficient facet search algorithms discussed heretofore, the lists LCF, LCF1, LCF2, and LNGF were all used in various ways to avoid attempting to solve each given problem Q more than once. In addition, the LCF (or the pair LCF1, LCF2) guided the sequence of facets investigated, and indicated when to stop investigating facets arising in the context of a particular efficient vertex. The latter functions are the "self fathoming" effect mentioned briefly in Chapter V. The price paid for all of these computational conveniences is the need for storage of extensive lists. The facet search algorithms discussed in this section do not require the storage of lists other than the LEV and LEF. On the other hand, a signal advantage of these lists--the guarantee against multiple attempts at the solution of any particular problem Q--is lost. The one essential function which the omitted lists provided is the task of guiding the search through the possible alternative subsets of the defining index set of each efficient vertex in such a manner that no efficient facets are overlooked. In the algorithms about to be presented, this function is furnished by a sequencing rule which requires negligible storage for its operation.

In the first algorithm here, sequencing is accomplished by associating each conceivable subset of the defining index set of the vertex in question with a nonnegative binary integer. Starting from zero, the integers are expressed in binary and this binary number is used to generate an index set. In this manner every possibility is

[illegible]

[illegible]

[illegible]

[illegible]

enumerated. (By contrast, the previous facet finders were able to avoid totally enumerating the index sets through capitalizing on the E-inference and NGE-inference theorems (V-14 and V-15) through the use of the LCF, or the LCF1 and LCF2.)

The binary number based sequencing rule operates as follows. To each index corresponding to an inequality constraint in the defining index set of the vertex under consideration is associated a column in the binary number system. Columns are assigned starting from the right; thus if there are four indices of inequality constraints in a given vertex's defining index set, then these would be identified with the "one's column", the "two's column", the "four's column" and the "eight's column". The defining index set associated with a given binary number contains two sets. The first is the set of indices of equality constraints. The second is the set of those inequality constraint indices whose columns in the binary number contain a numeral 1. For example, if $LEV_m = (S,i_1,i_2,i_3)$ and if i_1 is associated with the one's column, i_2 with the two's column, and i_3 with the four's column, then the index set associated with the integer 6--in binary "110"--is (S,i_3,i_2).

Each time a new vertex is taken under consideration, the sets associated with the integers starting from zero are considered; this LEV is complete when the integers get too large to be expressed in binary with the columns to which indices have been assigned. In the example just given, 7 is the largest integer possible (and even this is not necessary, for it is the entire set) because 8 requires

four binary columns, and no index has been assigned to the fourth--the eight's column.

A low storage algorithm based on binary sequencing:

Step #1: Store the defining index sets of the efficient vertices in $LEV_1, LEV_2, \ldots, LEV_U$. Set $i = 0$. Attempt to solve the problem Q(S): if a solution exists, place the set (S) on the LEF and stop. Otherwise go to step #2.

Step #2: Set $i = i+1$. If i exceeds U, stop. Otherwise set J" = the index set associated with LEF_i. Associate (from the right) columns in the binary number system with inequality constraint indices in J". Set $j = 1$.

Step #3: Ascertain whether the index set J associated with the number j expressed in binary contains an index set on the LEF. If not, attempt to solve the associated problem Q. If a solution exists, add J to the LEF and remove from the LEF any index set containing J.

Step #4: Set $j = j+1$. If j can not be expressed in binary with the columns to which indices are assigned, go to step #2. Otherwise go to step #3.

Example VI-3:

Consider the problem and notation established in example VI-1. (Recall that $LEV_1 = (S,1,2,3,4)$; LEF includes f(S,1,3) and f(S,4)). The low storage algorithm with binary sequencing proceeds

as shown in the following table:

Number ($Base_{10}$)	Binary Number: (i_1)	(i_2)	(i_3)	(i_4)	Set J	On the LEF?	Q(J)?	Notes
							No	Q(S)
1				1	(S,4)	No	Yes	to LEF
2			1	0	(S,3)	No	No	
3			1	1	(S,3,4)	Yes		
4		1	0	0	(S,2)	No	No	
5		1	0	1	(S,2,4)	Yes		
6		1	1	0	(S,2,3)	No	No	
7		1	1	1	(S,2,3,4)	Yes		
8	1	0	0	0	(S,1)	No	No	
9	1	0	0	1	(S,1,4)	Yes		
10	1	0	1	0	(S,1,3)	No	Yes	to LEF
11	1	0	1	1	(S,1,3,4)	Yes		
12	1	1	0	0	(S,1,2)	No	No	
13	1	1	0	1	(S,1,2,4)	Yes		
14	1	1	1	0	(S,1,2,3)	Yes		
15	1	1	1	1	(S,1,2,3,4)	Yes		
16								too large to be represented in 4 binary columns; go to LEV_2

Eight different problems Q were attempted.

[illegible]

Theorem VI-5: Solution x is efficient to problem P if and only if it is contained in some facet on the LEF generated by the algorithm just discussed.

Corresponding to any possible facet is a binary number. This number must be generated in step #2. Proof of the theorem follows directly from this, by arguments now familiar from the proofs of each of the preceding algorithms.

The following low storage algorithm employs a sequencing rule which provides the same sort of enumeration as is achieved by the algorithm of Chapter V. Since it omits the LCF of the earlier algorithm, it does not share that algorithm's efficient self fathoming feature. It does, however, have a crude self fathoming feature of its own based on the observation that if every conceivable facet whose defining index set has a given cardinality is efficient, then every facet of a larger cardinality must be efficient also.

A low storage facet finding algorithm with variable base sequencing:

Step #1: Store the defining index sets of the efficient vertices in LEV_i, $i = 1,2,\ldots,U$. Set $i = 0$. Attempt to solve the problem Q associated with set S: if a solution exists, place S on the LEF and stop. Otherwise set flag = 0.

Step #2: Set $i = i+1$. If i exceeds U, stop. Otherwise set J'' = the defining index set of LEV_i. Let n = the

number of indices in J"-S; associate each such index with an integer from 1 to n. Set j = 1.

Step #3: Express j in the base n+1 number system. Form the set J = S + every index whose associated numeral is in the base n+1 expression for j. Ascertain whether J contains any set on the LEF; if not, attempt problem Q(J), and if a solution exists, place J on the LEF. Otherwise set flag = 1.

Step #4: Set j = the next larger integer expressible in the base n+1 number system with the property that the numerals including zero in that expression are strictly decreasing from the right. If no such j exists, to step #2. If this j requires one more column in its base n+1 expression than the last, and if flag = 0, go to step #2. Otherwise set flag = 0 and go to step #3.

Example VI-4: Let LEV_1 = (S,1,2,3,4) and let the efficient facets containing this vertex be f(S,1), f(S,3), and f(S,2,4). The algorithm just discussed would proceed with this problem as shown in the following table. Notation is as introduced in example VI-1. "Skip" indicates that the base n+1 representations of the integers concerned do not consist of numerals which are strictly increasing from the left.

Number ($Base_{10}$)	Number ($Base_5$)	Set J	On the LEF?	Does Q(J) have a solution, if attempted?	Remarks
				No	Q(S)
1	1	S,1	No	Yes	to LEF
2	2	S,2	No	No	
3	3	S,3	No	Yes	to LEF
4	4	S,4	No	No	
5,6					skip
7	12	S,1,2	Yes		
8	13	S,1,3	Yes		
9	14	S,1,4	Yes		
10-12					skip
13	23	S,2,3	Yes		
14	24	S,2,4	No	Yes	to LEF
15-18					skip
19	34	S,3,4	Yes		All acceptable two digit $base_5$ numbers have been shown to have efficient sets J; proceed to LEV_2.

Ten sets have been enumerated and six problems Q have been attempted. Were the previous low storage algorithm to be applied, it would also attempt six problems Q, but would enumerate all sixteen possible sets.

[illegible]

The proof of the completeness of this algorithm is analagous to the proof of the completeness of the first low storage algorithm.

VI-5 COMMENTS AND SUMMARY

The author has no computational experience with the facet search algorithms of this chapter. Nevertheless, it is perhaps well to summarize the expected attributes of these approaches.

* The efficient facet search procedures of Chapter V and of section VI-2 are both straight forward in operation and reasonably simple to program for the computer. The choice between the two would depend most heavily on whether large (few tight constraints) or small (many tight constraints) facets are expected.
* The low storage procedures are the simplest algorithms of the lot, both conceptually, and in programming terms. In terms of the numbers of linear programs to be solved, they can generally be assumed to be inferior to the other approaches. This is because only the most rudimentary precaution against multiple solutions of a given problem Q is used. In addition, the number of comparisons required may occasionally be quite large, because every conceivable alternative subset may be generated and inspected.
* The double-ended facet finder has the greatest complexity and the greatest variety in its protections against duplication and solution of unnecessary linear programs. While it requires more lists and more computer program than the other approaches, its total storage requirements

might compare favorably with those of the Chapter V or the bottoms-up algorithms. Whether the number of problems Q attempted favors this algorithm, the bottoms-up algorithm, or the algorithm of Chapter V depends on the character of the problem at hand. Example problems favoring each have been found.

In summary, the author is not at this time able to demonstrate the clear superiority of any of the efficient manifold searches cited in this chapter and the last. Apparently there are countless other possible facet search procedures. For example, an interesting variation on those of this chapter would be a bottoms-up low storage approach. Moreover, any of these procedures can be modified in numerous ways (e.g., any one which employs the LNGF could choose to truncate it when it reaches a certain length). It would appear that the characteristics of all of these approaches are sufficiently diverse that any one of them could be at a disadvantage on some occasion. In choosing one for implementation (or in designing a new one), the analyst should be guided by the expected magnitude and complexity of the intended application, by the computer available, and perhaps even by the character of the linear programming packages available for solving the attendant subproblems.

CHAPTER VII

THE SCHOOL OF HEALTH SERVICES EQUILIBRIUM PLANNING MODEL: RESULTS

VII-1 A REVIEW OF THE EXPRESSION OF THE EFFICIENT MANIFOLD

Before discussing the efficient solutions to the problem of Chapter IV, the manner in which the set of efficient solutions is to be expressed will be briefly reviewed for the benefit of the reader who has skipped Chapters V and VI.

Imagine that it is desired to express all of the pairs (x_1,x_2) which, if plotted, would fall in the cross hatched region of Figure IV.

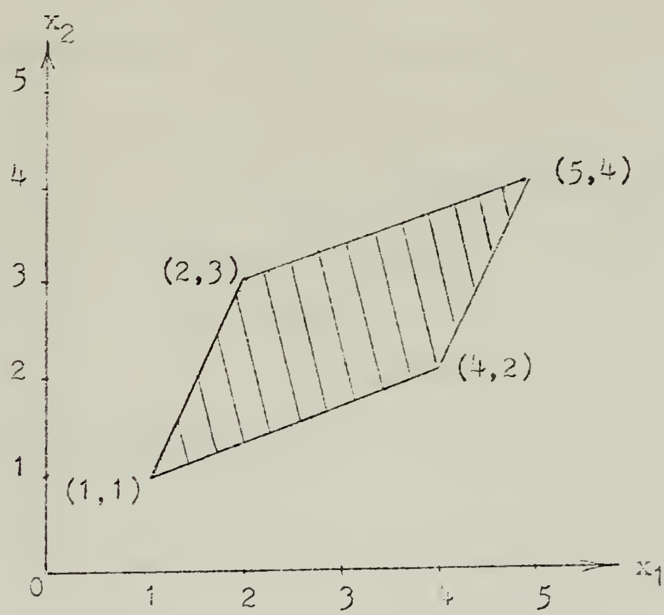


Figure IV

[illegible]

[illegible]

[illegible]
[illegible]
[illegible]
[illegible]
[illegible]

[illegible]
[illegible]
[illegible]

One way to accomplish this is to list the corners ("vertices") of the shaded area, remarking that these and any points which can be expressed as weighted sums of these, with nonnegative weights adding to 1, are in the shaded region. Such weighted sums are called "convex combinations". For example, in Figure V, the pair (3,2.5), which is labeled A, is a weighting of the vertices (2,3) and (4,2); in fact,

$$(3,2.5) = 1/2\ (2,3) + 1/2\ (4,2)$$

Similarly the point B, representing the pair (11/3,3) is

$$(11/3,3) = 1/3\ (5,4) + 1/3\ (2,3) + 1/3\ (4,2)$$

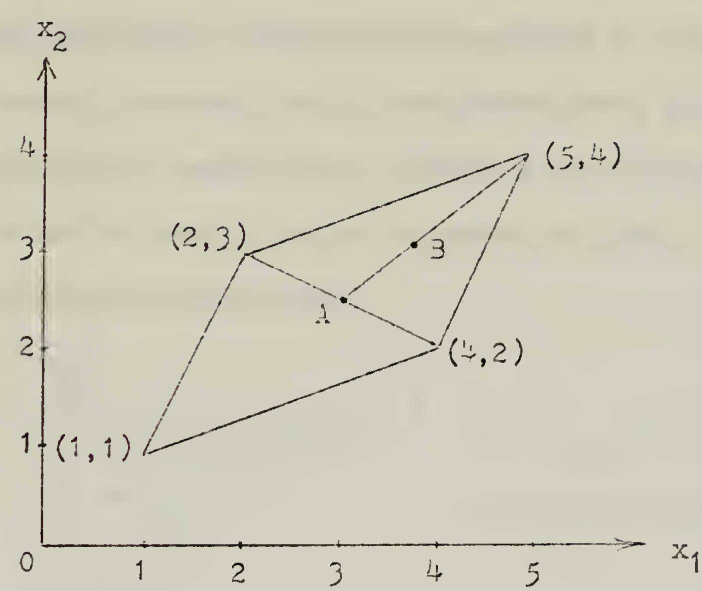


Figure V

The set of efficient solutions, that is the "efficient manifold", to the School of Health Services equilibrium planning model is expressed in this manner--by specifying its vertices. These particular results illustrate an occasion when the vertices must be organized into several sets. The totality of all convex combinations of the members of each such set is called a "facet". Members of different sets should not be used when forming convex combinations, for otherwise the result may not be efficient; i.e., may not fall within a facet. For example, Figure VI shows two (hatched) efficient facets together with the vertex sets generating them. Vertices B and D belong to both sets; that is, they are each contained in both facets. Incidentally, a consequence of the theory of Chapter V is that whenever an efficient manifold is constituted of more than one facet, then there will be enough vertices common to two or more facets that the efficient manifold is connected, as this one is. Vertex A belongs only to the first facet and vertex B belongs only to the second facet. Point E is a convex combination of A and C; notice that it is not a member of either efficient facet and is therefore not efficient.

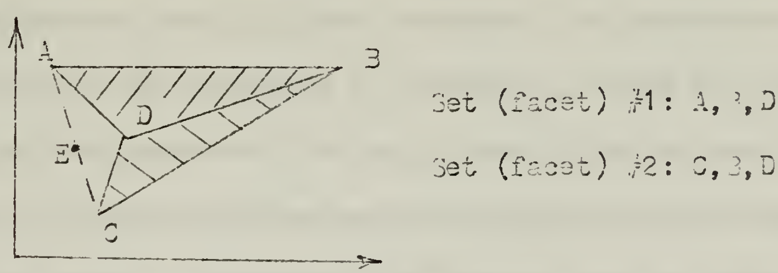


Figure VI

VII-2 THE EFFICIENT MANIFOLD

When the efficient manifold finding technique of Chapter V is applied to the problem expressed in section IV-4, 108 efficient vertices are found. These are arrayed in two efficient facets. Facet #1, containing 88 efficient vertices, is the set of all efficient solutions having additional clinical supervisor availability as their common limiting constraint. Facet #2, containing 52 efficient vertices, is the set of all efficient solutions having faculty office and dry laboratory availability as their common limiting constraints. It happens that 32 vertices are common to both facets.

One way to characterize the vertices is by the values which the problem variables assume at them. This is what was done when the 108 count just mentioned was found. The model contains 17 variables including slack variables (a slack variable is a quantity representing the amount of underutilization of the resource on the right hand side of one of the "less than or equal to" constraints). It is also possible to characterize the vertices according to the criteria values achieved at them; this is the natural way for the decision maker to view them. There are only five criteria, and these are determined from only four variables. In particular, the slack veriables and the variables Y_1 through Y_4, which are concerned with office and classroom utilization, do not figure directly in the annual numbers of graduates from each curriculum or in the annual surplus of the school. Consequently efficient vertices which differ only in their values of these variables are, from the standpoint of

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

the criteria, identical. This phenomenon is so widespread in these particular results that among the 108 different efficient vertices only fifteen different criteria scores patterns are represented. Nine of these patterns belong to Facet #1, including four which also belong to Facet #2. Including the latter four, there are ten patterns in the second facet. Table I is a listing of these patterns according to whether they belong only to the first facet (Group A), to both facets (Group B), or only to the second facet (Group C). The values of the first four criteria have been rounded off to the nearest integer. These are, of course, the annual numbers of graduates from the programs indicated. Since it has been assumed that there will be no dropouts, and since each curriculum has a two year duration, these figures may also be considered to represent the annual matriculations or half of the annual enrollments in the curricula represented.

The reader is reminded that the surpluses shown in Table I, the measurements according to the fifth criterion, are <u>variable</u> instructional surplus. That is, only those costs and incomes which are judged to be proportional to enrollments or faculty size are included. In order to find the total surplus corresponding to any of these it would be necessary to add to the figure shown another figure representing the sum of all other incomes such as grants unrelated to the enrollments of the School, minus the sum of all other expenses, such as the salaries of the faculty specialists. But notice that the <u>relative</u> impact of two different enrollment and

(Vertex label)	Health Assistants	H. Assoc./ N. Prac.	H. Serv. Mgmt.	Env. Hygiene	Instructional Surplus ($000)
Group A:	Vertices in facet #1 alone:				
A	69				
B	69		2		-7
C	68			7	-34
D	63	6			-77
E		69			-821
Group B:	Vertices in facet #1 and facet #2:				
F	61		97		-449
G	47			122	-579
H		61	86		-1,128
I		48		111	-1,111
Group C:	Vertices in facet #2 alone:				
J	47		107		-495
K	23			143	-675
L		34	110		-920
M		16		145	-884
N			141		-657
P				162	-769

TABLE I

	Minimum	Maximum
Annual graduates:		
Total	69	169
H. Assistant program	0	69
H. Associate and Nurse Practitioner programs	0	69
H. Services Mgmt. program	0	141
Environmental Hygiene program	0	162
Instructional Surplus ($000)	$-1,128	$0

TABLE II

staffing programs can be obtained simply by comparing the surpluses shown in the table. For example, pattern I, with a variable surplus of $-1,111,000 is $662,000 more expensive per year than pattern F, with a variable surplus of $-449,000, regardless of the magnitude of the non-proportional incomes and expenses.

It happens that the maxima and minima of the values attained by the criteria over the set of all efficient solutions are achieved at vertices. The maxima and minima for the criteria of this problem and the maximum and minimum values of the sum of the annual graduations, or matriculations, are shown in Table II. These figures indicate the approximate dimension of the efficient manifold.

Any convex combination of any of the patterns in groups A and B in Table I corresponds to an efficient solution in Facet #1. Similarly, any such combination of some or all of the patterns in groups B and C is an efficient solution in Facet #2. Together these constitute the set of all efficient possibilities. Care should be taken to avoid mixing patterns from the first and third groups; such combinations are not necessarily efficient.

The following examples illustrate the use of Table I to generate additional efficient matriculation patterns. Consider first a straightforward weighted averaging of members of Facet #1. Imagine that a decision maker has reviewed groups A and B and decides to investigate the average of patterns C and H. He calculates as follows:

$$\begin{array}{lll} & 1/2 \text{ of } (68,\ 0,\ 0,\ 7,\ -34) & = 1/2 \text{ of pattern C} \\ + & \underline{1/2 \text{ of } (\ 0,61,86,\ 0,-1128)} & = 1/2 \text{ of pattern H} \\ = & \quad (34,31,43,\ 4,\ -581) & \end{array}$$

That is, an equal averaging of these patterns yields an efficient solution in which 34 health assistants, 31 health associates or nurse practitioners, 43 health services managers, and 4 environmental hygienists are matriculated and graduated annually, and which has an annual variable instructional budget cost of $581,000. Perhaps on examining this solution, the decision maker decides that he would prefer more environmental hygienists; he might add a dash of pattern G, as follows:

$$\begin{array}{lll} & .8 \text{ of } (34,31,43,\ \ 4,\ -581) & = .8 \text{ of the previous result} \\ + & \underline{.2 \text{ of } (47,\ 0,\ 0,122,\ -579)} & = .2 \text{ of pattern G} \\ = & \quad (37,25,34,28,\ \ -581) & \end{array}$$

and so on.

One may also experiment graphically with these possibilities. In Figure VII, the criteria levels of patterns J and P from group C in Table I are marked on the left and right axes respectively. Notice that there are two scales shown; one for the numbers of graduates of the various programs, and the other for the annual variable surplus. Corresponding criteria for the two patterns have been joined by straight lines and these are labeled to show which

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

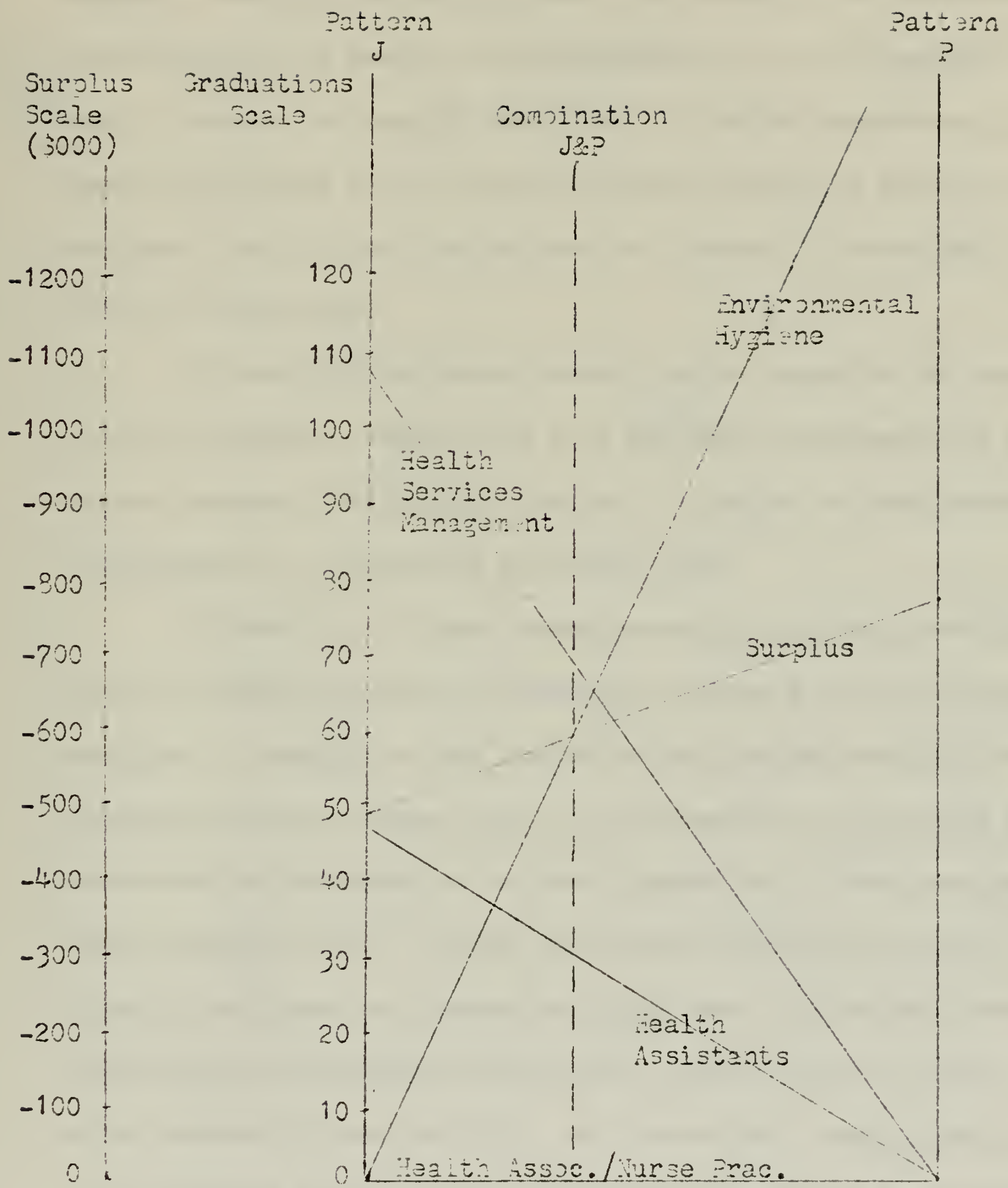

Surplus Scale ($000)
Graduations Scale
Pattern J
Combination J&P
Pattern P
-1200
-1100
-1000
-900
-800
-700
-600
-500
-400
-300
-200
-100
0
120
110
100
90
80
70
60
50
40
30
20
10
0
Environmental Hygiene
Health Services Management
Surplus
Health Assistants
Health Assoc./Nurse Prac.

FIGURE VII

criteria they represent. The intersections of these links with any vertical line drawn between the two vertical axes correspond to the scores attained by some combination of the solutions J and P. For example, reading from the top down, the dashed line represents a pattern having 70 health services managers, 60 environmental hygienists, a surplus of roughly $-580,000, 30 health assistants, and no health associates or nurse practitioners graduating annually. Like patterns J and P, this new pattern is a member of Facet #2; therefore it is efficient.

If the decision maker wishes, he may transfer the results of one such graphical combination to a new axis, complement it with yet another pattern, and combine the two. A series of such successive combinations is illustrated in Figure VIII.

In addition to these composite matriculations patterns, there is other interesting information provided by the efficient manifold. A perusal of the values of the problem variables at the efficient vertices reveals that no case exhausts the supply of demonstration laboratories or bench laboratories which was postulated in Chapter IV. In fact, the lowest surplus capacities reported for these two laboratory types were 8 hours per week of demonstration laboratory availability, and 13 hours per week of bench laboratory availability. The reader will recall that no reliable estimates of the true availability of bench or demonstration laboratory space was available, but that a reasonable minimum of one full time laboratory of each type was assumed. These results show

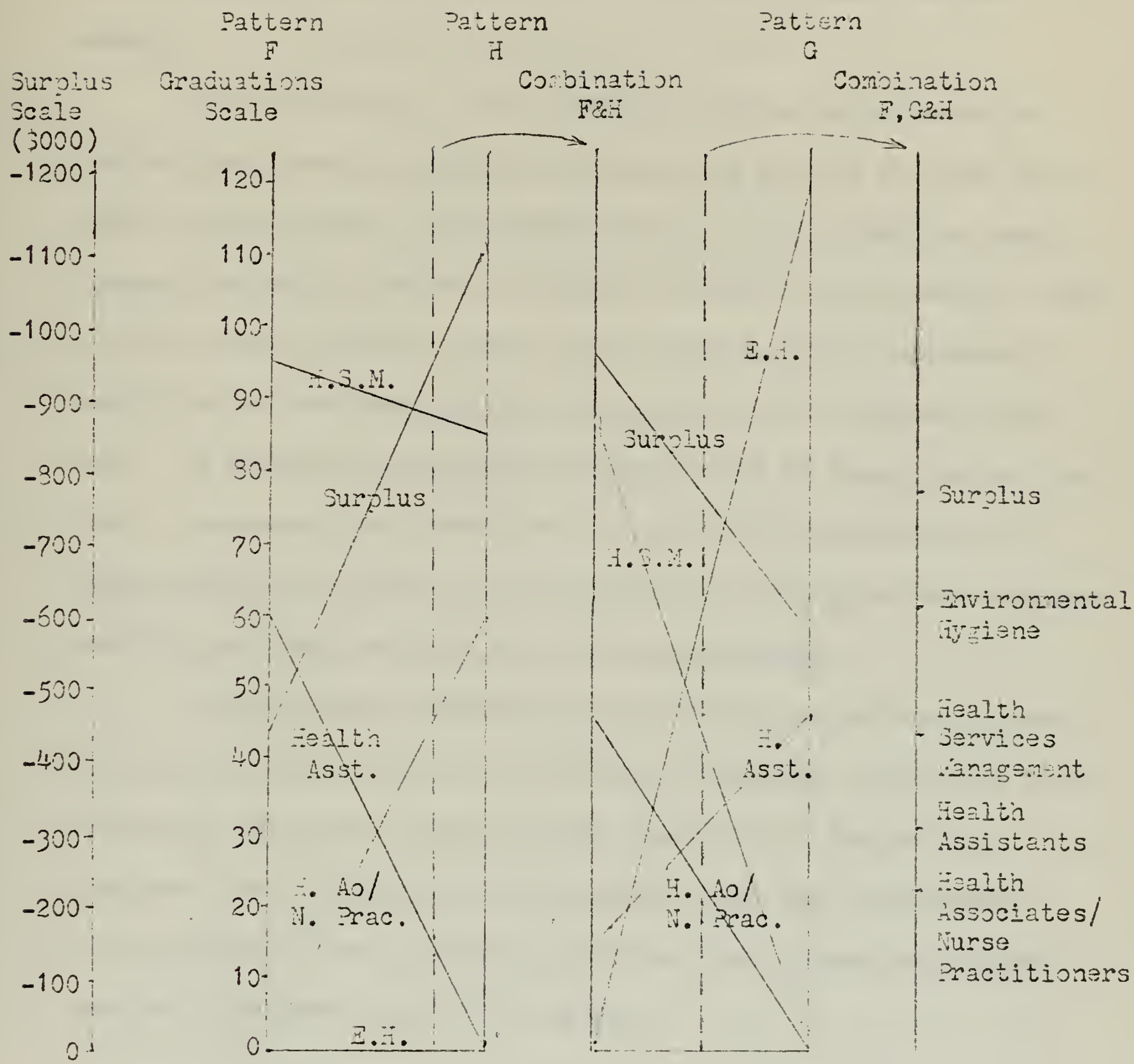

Pattern F
Graduations Scale
Pattern H
Combination F&H
Pattern G
Combination F,G&H
Surplus Scale ($000)
-1200
-1100
-1000
-900
-800
-700
-600
-500
-400
-300
-200
-100
0
120
110
100
90
80
70
60
50
40
30
20
10
0
H.S.M.
Surplus
Health Asst.
H. Ao/ N. Prac.
E.H.
Surplus
H.S.M.
E.H.
H. Asst.
H. Ao/ N. Prac.
Surplus
Environmental Hygiene
Health Services Management
Health Assistants
Health Associates/ Nurse Practitioners

FIGURE VIII

that if only this bare minimum were to be provided, the demonstration laboratory would stand empty at least one day weekly, and the bench laboratory would be empty, at a minimum, almost two days weekly.

The variable Y_4, which represents the amount of usage of medium classrooms for undersized classes, is present at every efficient vertex; however, its minimum value is only 4; that is, small classes are held in medium classrooms at least 4 hours weekly. When a small class occupies a large room it does so quite comfortably, regardless of how much too large the room is. The reverse is not true. In allocating available classroom space to large, medium, and small classrooms, the greater evil is to err in the direction of rooms which are too small. This positive Y_4 throughout the efficient manifold confirms that this error has been avoided.

It is perhaps interesting to note the computational characteristics of this solution. The author's computer code, which is an inexpertly programmed example of the algorithm of Chapter V, required 8 and 1/2 minutes for execution on an IBM 7094 computer. 1702 problems "Q" were attempted, and the list of candidate facets attained a maximum length of 74 facets.

CHAPTER VIII

CONCLUSIONS AND RECOMMENDATIONS

VIII-1 THE EFFICIENT MANIFOLD FINDING PROCEDURE

The problem addressed in this thesis is to find the set of efficient solutions to

$$\begin{aligned} &(\text{Maximize}) \quad F(x) \\ &\text{subject to} \quad G(x) \leqq 0 \end{aligned}$$

where F(x) and G(x) are linear vector valued functions of the vector x, a member of Euclidean n-space. A straightforward, finite procedure for doing so has been revealed, and several variations on the basic method have been presented. The method and its several versions all rely in principle upon ordinary linear programming notions and in execution upon Simplex algorithm procedures. The efficient manifold finding procedure has been illustrated by application to a problem of planning in higher education, and the characteristics expected of the alternate facet finding procedures have been noted.

The successful application of the efficient manifold finding procedure of Chapter V to the model of Chapter IV demonstrates that

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

the algorithm does indeed perform as anticipated. In Chapter VII it was established by example that the efficient manifold so presented can indeed provide helpful information. This success notwithstanding, certain improvements in the approach may be anticipated.

The reader may have noted that the efficient manifold of Chapter VII is particularly simple. Much more involved results will occasionally arise from problems nearly as small, and quite moderate problems may have efficient manifolds which are beyond the capability of the decision maker to handle when presented in the artless manner illustrated. Clearly, better styles of presentation and better ways to cope with involved efficient manifolds are in order.

Quite a number of variations on the basic efficient manifold finding procedure were suggested in Chapter VI. Exploration of these and other possibilities should continue; considerable streamlining may be possible. Two intriguing possibilities not previously mentioned are the following. First, perhaps some sort of accelerator might be developed for use at degenerate vertices. The number of bases associated with a degenerate vertex grows monstrously as the number of excess zeros increases; depending on the circumstances, many or all of these may meet the efficiency criteria. The algorithms presented must investigate them all, just as if they represented distinct vertices; perhaps this could be avoided. Secondly, a more complete merger of the vertex finding and facet finding phases of the approach might prove advantageous. Such an approach might involve making all calculations from a single tableau, a sort

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

of extended "problem Q" tableau (see Chapter V) in which every constraint gradient is present but of which certain such gradients are ineligible for introduction into the basis in any given problem. It will be recalled that a vertex is a facet of cardinality m--which suggests a way to find an efficient, though not necessarily feasible, vertex directly from this tableau.

One of the particularly interesting and useful capabilities of goal programming is its ability to optimize several objectives serially--e.g., a second criterion might be optimized subject to the condition that the optimality of the first criterion is maintained. Such an option would be a welcome addition to the techniques of the efficient manifold presentation method. The idea is to be able to eliminate a variety of multiple optimality which may arise when some variables are not represented in the natural criteria. Such a case arose in the model of Chapter IV: a mere 15 configurations of the variables of interest were represented by 108 distinct efficient vertices. It is possible to introduce an essentially arbitrary criterion in which the extra variables are featured, but if this new criterion is accorded equal status with the others then in addition to eliminating some redundancy it may also introduce new vertices which are efficient in the sense that some or all of the original criteria may be improved at the expense only of this secondary measure. Clearly these new solutions are quite as objectionable as the redundant ones which were eliminated. A preemptive priority procedure along the lines of the one used in goal programming might

perhaps be used to produce the desired effect without introducing these unwanted vertices.

Finally, there may be other good ways to go about finding efficient manifolds. It is possible, for example, that some benefit may be had from the knowledge of which vertices neighbor others. This information is immediately available from the lists of variables basic at each given vertex.

VIII-2 APPLICATIONS TO PLANNING IN HIGHER EDUCATION

The last chapter reported on the application of the method of this thesis to a model of the School of Health Services. The Dean of that School found this small study to be rather enlightening; during the formulation of the model, certain previously unappreciated aspects of the School's operations were deservedly highlighted. Chief among these was the significance of the supply and expense of part-time clinical supervisors. The nature of the efficient manifold which was revealed seems to confirm the importance of the availability of these particular faculty. In addition to providing early identification of the presently existing limitations to the School's growth, the efficient manifold is expected to be of value to the staff of the School in preparing funding requests.

In recent years the demand for analytic tools for institutional planning has markedly increased. This trend is a response to many pressures, including a sharpened appetite on the part of the state and federal governments for quantitative documentation of needs and performance. The chief impediment has been a vexing inability to satisfactorily treat with the outputs of higher education. A great deal of attention has been paid to this area--one bibliography lists almost two hundred references[1]--but the problems remain. A major difficulty seems to be that the products of education enjoy an ample inventory of genuinely incommensurable aspects. Impending developments and recent work in multicriterion problem solving may yet enable institutional planners to cope with this

[illegible]

[illegible]

troublesome diversity. The author believes that the efficient manifold presentation method holds particular promise because of its generous accommodation to the difficult circumstances of the decision maker.

FOOTNOTES

[1]Powel, John H. and Robert D. Lamson, An Annotated Bibliography of Literature Relating to the Costs and Benefits of Graduate Education, The Council of Graduate Schools, Washington, D.C., 1972.

APPENDIX I

In section V-3, following theorem V-6, it was observed that not all bases for a degenerate efficient vertex need have the property that for their respective reduced costs matrices G there exists vector $L > 0$ such that $LG \leqq 0$. The following example illustrates such a case.

Consider the following problem:

$$(\text{Max}) \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$\text{Subject to} \begin{bmatrix} 1 & 1 & 1/2 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \leqq \begin{bmatrix} 2 \\ 1 \\ 1 \\ 1 \end{bmatrix} \quad \begin{array}{c} \underline{\text{Slack Variables}} \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{array}$$

$$x_1, x_2, x_3 \geqq 0$$

The solution $x' = (1,1,0,0,0,0,1)^t$ is a degenerate vertex. Consider the basis B comprised of the columns corresponding to x_4, x_1, x_2, and x_7; from this, the reduced costs matrix G is developed, as follows:

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

$$B = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$B^{-1} = \begin{bmatrix} 1 & -1 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$C^B = \begin{array}{c} \begin{array}{cccc} x_4 & x_1 & x_2 & x_7 \end{array} \\ \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

$$C^B B^{-1} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\{Z_j\} = C^B B^{-1} a_j = \begin{array}{c} \begin{array}{ccc} x_3 & x_5 & x_6 \end{array} \\ \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \end{array}$$

$$\{C_j\} = \begin{array}{c} \begin{array}{ccc} x_3 & x_5 & x_6 \end{array} \\ \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \end{array}$$

$$G = \{C_j\} - \{Z_j\} = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ 1 & 0 & 0 \end{bmatrix}$$

Clearly there is no $L > 0$ such that $LG \leqq 0$. Nevertheless this vertex is efficient, as will now be demonstrated. The constraints exactly satisfied by x^1 are the first, second, and third, and the nonnegativity constraint for x_3, the relevant form of which is $-x_3 \leqq 0$. The matrix A of the gradients of these constraints is

$$A = \begin{bmatrix} 1 & 1 & 1/2 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

For vectors $k = (1,0,0,0)$, which is nonnegative, and $L = (1,1,1/2)$, which is strictly positive, the condition $kA = LC$ of theorem V-3 is met; the point is efficient.

APPENDIX II

In section V-5 it was remarked that a facet every extreme point of which is efficient is not itself necessarily efficient. The following example illustrates this point. Consider the problem

$$(\text{Max}) \quad \begin{bmatrix} -10 & 1 & 1 \\ 1 & -10 & 1 \\ 1 & 1 & -10 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$\text{Subject to} \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \leqq 1 \quad (\text{Slack variable: } x_4)$$

$$x_1, x_2, x_3 \geqq 0$$

This problem is graphically represented in the following figure. The feasible region is the tetrahedron in the positive orthant.

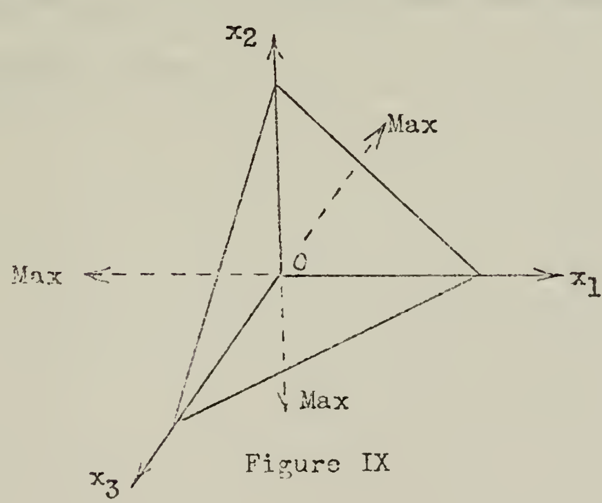
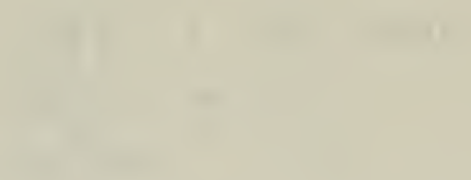


Figure IX

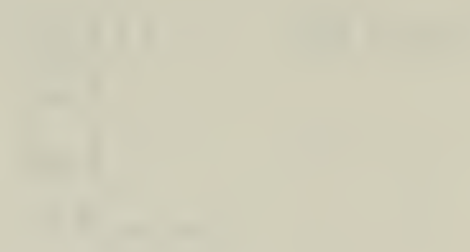
[illegible]

[illegible]

[illegible]



[illegible]



[illegible]

Consider the facet $f(x_4)$, the set of all feasible solutions for which the first constraint is tight. In the figure, this is the triangular face toward the veiwer. This facet has three vertices; they are

$$x^1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad x^2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \qquad x^3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

Each of these will be shown to be efficient. At x^1, the first constraint and the nonnegativity constraints for x_2 and x_3 are tight; then the matrix of coefficients of tight constraints A is

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

Let $k = (1,0,99)$ and $L = (1,1,10)$; these are nonnegative and positive respectively. Then

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

$$kA = (1,0,99)\begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} = (1,1,-98)$$

$$= (1,1,10)\begin{bmatrix} -10 & 1 & 1 \\ 1 & -10 & 1 \\ 1 & 1 & -10 \end{bmatrix} = LC$$

and, by theorem V-3, x^1 is efficient. Similarly, at x^2,

$$A = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

and letting $k = (1,0,99)$ and $L = (1,1,10)$ we have $kA = LC$.

Finally, at x^3,

$$A = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \end{bmatrix}$$

Letting $k = (1,0,99)$ and $L = (1,10,1)$, $kA = LC$.

Nevertheless, $f(x\ )$ is not efficient, for there are no $k \geqq 0$, $L > 0$ such that

$$k\ (1,1,1) = L\begin{bmatrix} -10 & 1 & 1 \\ 1 & -10 & 1 \\ 1 & 1 & -10 \end{bmatrix}$$

[illegible]

[illegible]

$$[illegible]$$

[illegible]

$$[illegible]$$

[illegible]

$$[illegible]$$

[illegible]

$$[illegible]$$

Thus, for example, the feasible solution $x^4 = (1/3,1/3,1/3,0)$, which is in $f(x_4)$ and satisfies no constraint other than that defining this facet, is not efficient.

[illegible]

[illegible]

[illegible]

BIBLIOGRAPHY

Balderston, F. E. and G. B. Weathersby, PPBS in Higher Education Planning and Management from PPBS to Policy Analysis, University of California, Research Program in University Administration, paper P-31, May 1972.

Benayoun, R. and J. Tergny, "Critères multiples en programmation mathematique: une solution dans le cas linéaire." Revue Française d'Informatique et de Recherche Operationnelle, 3ème année, no. V-2 (1969), pp. 31-56.

Charnes, A. "Duality and Degeneracy in Linear Programming," Econometrica, XX (1952), pp. 160-170.

Charnes, A. and W. Cooper, Management Models and Industrial Applications of Linear Programming, Vol. 1, John Wiley and Sons, New York, 1961.

Dyer, J. S. "A Time-Sharing Computer Program for the Solution of the Multiple Criteria Problem," Management Science, XIX (1973), Applications series, pp. 1379-1383.

Dyer, J. S. "Interactive Goal Programming," Management Science, XIX (1972), pp. 62-70.

Freimer, M. and P. L. Yu, An Approach toward Decision Problems with Multiobjectives, University of Rochester, Center for System Science, paper CSS 72-03, June 1972.

Geoffrion, A. M. "Proper Efficiency and the Theory of Vector Maximization," Journal of Mathematical Analysis and Applications, 22 (1968) pp. 618-630.

Geoffrion, A. M., J. S. Dyer, and A. Feinberg, "An Iterative Approach for Multi-Criterion Optimization with an Application to the Operation of an Academic Department," Management Science, XIX (1972), pp. 357-368.

Hill, J. S. and R. C. Judd, Finding Analytic Meaning in Enrollment Matrices, University of Toledo, Office of Institutional Research, 1972.

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

Ijiri, Y. Management Goals and Accounting for Control, North-Holland Publishing Company, Amsterdam, 1965.

Judy, R. W., J. B. Levine, and S. I. Center, CAMPUS V Documentation, Vols. 1-6, Systems Research Group, Toronto, Canada, 1970.

Lee, S. M. and E. R. Clayton, "A Goal Programming Model for Academic Resource Allocation," Management Science, XVIII (1972) Applications series, pp. B395-B408.

Lee, S. M. and W. Sevebick, "An Aggregation Model for Municipal Economic Planning," Policy Sciences, II (1971), pp. 99-115.

McNamara, J. F. "Mathematical Programming Applications in Educational Planning," Socio-Economic Planning Sciences, VII (1973), pp. 19-35.

Mangasarian, O. L. Nonlinear Programming, McGraw-Hill, New York, 1969.

Miller, D. W. and M. K. Starr, The Structure of Human Decisions, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1967 (page 56).

Mintyberg, H. "Managerial Work: Analysis from Observation," Management Science, XVIII (1971), Applications series, pp. B97-B110.

Philip, J. "Algorithms for the Vector Maximization Problem," Mathematical Programming, II (1972), pp. 207-229.

Powel, J. H., Jr., and R. D. Lamson, Elements Related to the Determination of Costs and Benefits of Graduate Education, The Council of Graduate Schools, Washington, D.C., March 1972.

__________, An Annotated Bibliography of Literature Relating to the Costs and Benefits of Graduate Education, The Council of Graduate Schools, Washington, D.C., March 1972.

Resource Requirements Prediction Model (RRPM) 1.6; Users Manual, National Center for Higher Education Management Systems at WICHE, Boulder, Colorado, 1972.

Roy, B. "Problems and Methods with Multiple Objective Functions," Mathematical Programming, VI (1971), pp. 239-266.

Saska, J. "Linear Multiprogramming," Economiko Matematiky Obzor, IV (1968), pp. 359-373.

Schroeder, R. G. "A Survey of Management Science in University Operations," Management Science, XIX (1973), Applications series, pp. 895-906.

Simonnard, M. Linear Programming, Prentice-Hall, Englewood Cliffs, New Jersey, 1966.

Smith, G. T. and N. A. Baxter, University of Nevada Medical Education Feasibility Study, Nevada State Bureau of Business and Economic Research, November 1968.

Vazsonyi, A. "Why Should the Management Scientist Grapple with Information Systems," Interfaces, III, No. 2, pp. 1-18.

Weathersby, G. B. and M. C. Weinstein, A Structural Comparison of Analytical Models for University Planning, University of California, Research Program in University Administration, paper P-12, August 1970.

Yu, P. L. Cone Convexity, Cone Extreme Points and Nondominated Solutions in Decision Problems with Multiobjectives, University of Rochester, Center for System Science paper CSS 72-02, April 1972.

VITA

Stephen Trygve Holl was born in Minneapolis, Minnesota on the tenth of May 1945. He attended various elementary and secondary schools and entered the United States Naval Academy in 1963. His final year at the Naval Academy was spent largely in independent research as a Trident Scholar, studying a problem of antisubmarine warfare. He graduated with distinction in 1967 and upon receiving the Bachelor of Science was commissioned an Ensign in the United States Navy. For two years thereafter he served in the engineering department in USS *Hollister* (DD-788). In 1969, under scholarship from the U.S. Navy's Junior Line Officer Advanced Scientific Educational Program, he began graduate study at the Johns Hopkins University in the department of Operations Research and Industrial Engineering. In 1973, upon completion of the requirements for the Ph.D., he will enter Destroyer School at Newport, Rhode Island, and from there he will return to sea.